Theoretical Studies of Interactions in Ionic Molecular Crystals

A thesis submitted in fulfilment of requirements
for the degree of Doctor of Philosophy

Nicole Ann Benedek

Bachelor of Applied Science (Applied Chemistry)(Honours)

School of Applied Sciences
Science, Engineering and Technology Portfolio
RMIT University
April 2006

Szeretettel a szüleimnek

Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of this thesis is the result of work which has been carried out since the official commencement date of the approved research program; and, any editorial work, paid or unpaid, carried out by a third party is acknowledged.

Nicole Ann Benedek

April 2006

Acknowledgements

I firstly wish to thank my supervisors, Dr. Kay Latham and Prof. Irene Yarovsky, for teaching me not only how to do good science but also how to be a good scientist. I would also like to thank Prof. Ian Snook, for many helpful discussions and assistance, particularly concerning QMC. I am especially grateful to all three for the support I received on two trips abroad.

My sincere thanks to Prof. Richard Needs and Dr. Mike Towler, who welcomed me to their group in the Cavendish Laboratory on two occasions and who have patiently assisted me in my QMC research since I started my PhD. Thanks are also due to Arash, Victoria, Matt and Nick for providing the fun and laughs which made my visits so memorable.

Thank you to all the members of the RMIT Outdoors Club for dragging me out of the office when I needed it most, in particular: Patrick (for geeky tech-talk), Garry (ditto geeky tech-talk), Darryl, Derek, Sven, Chris and Katherine. I am particularly indebted to Thalia and Gerard. My thanks also to The Girls – Zoe, Jaime and ZT – for their friendship and support. Britta Drevermann of the Chemistry department at RMIT has also been a great friend. The members of the RMIT Sims Group (especially Rob, Akin, Ryan, Michelle, David and Ish) have made it a pleasure to come to work everyday.

Special thanks to James, who unflinchingly supported me through some of my darkest hours. Words can't express...

This thesis is for my family, who patiently tolerated the mumblings, grumblings and bad moods which occasionally afflict PhD students and whose love and support made this work possible.

Finally, this project made extensive use of the computational resources provided by the Australian Partnership for Advanced Computing (APAC) and Victorian Partnership for Advanced Computing (VPAC).

Refereed Publications

1. N. A. Benedek, I. K. Snook, M. D. Towler and R. J. Needs, "Quantum Monte Carlo Calculations of the Dissociation Energy of Water Dimer" *J. Chem. Phys.*, **In press** (2006).

2. N. A. Benedek, K. Latham, I. K. Snook and I. Yarovsky, "Density Functional Theory Study of Hydrogen Bonding in Ionic Molecular Materials" *J. Phys. Chem. B*, **In press** (2006).

3. N. A. Benedek, I. K. Snook, K. Latham and I. Yarovsky, "Application of Numerical Basis Sets to Hydrogen Bonded Systems: A Density Functional Theory Study" *J. Chem. Phys.*, **122**, 144102 (2005).

4. N. A. Benedek, I. Yarovsky, K. Latham and I. K. Snook, "Quantum Monte Carlo Study of Water Molecule: A Preliminary Investigation" *Aust. J. Chem.* **57**, 1229 (2004).

5. N. A. Benedek, M. J. S. Spencer, K. Latham and I. Yarovsky, "Hydrogen Bonding in Mixed Ligand Copper Organophosphonates" *Chem. Phys. Lett.* **378**, 400 (2003).

Summary

We have used *ab initio* computational simulation techniques to investigate both intra- and intermolecular interactions in a novel family of ionic organophosphonate molecular crystals. We have examined the influence of various numerical approximations on the computed geometry and binding energies of a selection of well-characterised hydrogen bonded systems. It was found that numerical basis sets provided the efficiency required to study the large hydrogen bonded dimer anions present in the organophosphonate system, while also producing accurate geometries and binding energies. We then calculated the relaxed structures and binding energies of phenylphosphonic acid dimer in the two arrangements in which it is present in the bulk crystal. The computed geometries were in excellent agreement with the experimental structures and the binding energies were consistent with those found for other ionic hydrogen bonded systems. Electron density maps were used to gain insight into the nature of the hydrogen bonding interaction between phenylphosphonic acid dimers. We also examined the effect of aromatic ring substituents on the geometry and energetics of the hydrogen bonding interaction. The nitro-substituted dimer was predicted to have a stronger binding energy than its unsubstituted parent while the methyl-substituted dimer was predicted to have a similar binding energy to its unsubstituted parent. An analysis of crystal field effects showed that the structure of the phenylphosphonic acid dimers in the organophosphonates is a complex product of competing intra- and intermolecular forces and crystal field effects. Cooperative effects in the organophosphonate system were also investigated and it was found that the interactions were mostly one-body (local) in nature.

We have examined the intramolecular charge-transfer interaction between copper-halogen cations in the organophosphonate materials. The origin of geometric differences between the Cu(I) starting material and Cu(II) product cations was attributed to the electronic configuration of the Cu ion, not crystal field effects. To gain further insight into the difference in electronic structure between the starting material and product, we attempted to simulate the step-by-step dissociation of the $[CuI]^+$ system. Although this investigation was not successful, we were able to expose some of the pitfalls of simulating dissociating odd-electron systems. We also analysed and compared

the charge-transfer interaction in the chloro-, bromo- and iodo-forms of the organophosphonate family. The charge-transfer interaction was predicted to increase on going from the chloro- to the iodo-form, consistent with solid-state UV-visible data.

Finally, we used the highly accurate Quantum Monte Carlo (QMC) method to investigate the hydrogen bonding interaction in water dimer and to calculate the dissociation energy. The accuracy of the experimental estimate for the dissociation energy has recently been questioned and an alternative value has been put forward. Our results lend support to the validity of the alternative value and are also in excellent agreement with those from other high-level calculations. Our results also indicate that QMC techniques are a promising alternative to traditional wavefunction techniques in situations where both high accuracy and efficiency are important.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Molecular Materials

The discovery of electrically conductive polymers [5, 6] in the early 1970s sparked an explosion of interest in the properties of molecular materials. Semiconducting materials consisting of small organic molecules are promising candidates for optoelectronic devices and have thus been the subject of numerous theoretical [7, 8, 9] and experimental [10, 11] studies. The properties of molecular solids are dictated by features of both molecular and condensed matter physics: their physical properties depend on their crystalline structure, which in turn is governed by interactions between molecules. An understanding of the relationship between intermolecular interactions and electronic and optical properties is thus vital to the efficient design of new devices.

In contrast to their more conventional organic counterparts, inorganic molecular materials may contain any element. A far more diverse range of intra- and intermolecular interactions is therefore possible. Many inorganic molecules are charged, and hence also have counter-ions in condensed phases. This has consequences for both the diversity of interactions, the energies associated with them and the material properties. [12] Inorganic molecular solids have been extensively studied in the context of supramolecular science, a branch of chemistry concerned with the controlled assembly of complex molecules into molecular solids. [13, 14, 15] The ultimate aim of the supramolecular

chemist is to design materials with specific properties, a task that requires manipulation of delicate intermolecular interactions. Here, as with organic solids, the relationship between intermolecular interactions and the properties of interest must be understood.

The organophosphonates are an interesting class of molecular materials which typically contain both organic and inorganic components. [16] The structures of organic phosphates and phosphonates often resemble those of their inorganic analogues. In contrast, metal organophosphonates can exhibit a wide range of structures including monomers, one-dimensional chains, two-dimensional layers and three-dimensional pillared frameworks. [17] Research on these materials began in the 1950s, when it was discovered that layered metal phosphates could be used as cation exchangers in radioactive waste streams. [18] Phosphonates have since found applications as sorbents, catalysts [19] and gas sensors [20] and thus continue to enjoy the significant interest of researchers.

An interesting family of novel ionic molecular crystals, mixed-ligand copper organophosphonates, was recently synthesised and characterised. [3] These layered materials contain a copper ion (Cu^{2+}, d^9) co-ordinated to two phenanthroline ligands and a halogen substituent, either I^-, Br^-, Cl^- or NCS^-. The positively charged copper-halogen ions are interleaved between parallel sheets of negatively charged, hydrogen-bonded phenylphosphonic acid dimers, which act as counter-ions. There appear to be no covalent bonds between individual molecules, either within a layer or between layers. The compounds crystallise in a monoclinic space group, C2/c, and have their Cu-X bond lying along a two-fold crystallographic axis. The molecular unit of the iodo-form is shown in Figure 1.1 and a 1×2×1 supercell is shown in Figure 1.2, where 1×2×1 means one conventional cell length in the a direction, 2 cells lengths in the b direction and one cell length in the c direction.

Our initial interest in these materials stemmed from the observation of some curious properties of the iodo-form, which were different to those of the other complexes. Initial measurements of the bulk magnetic properties of the iodo-species suggested the presence of a diamagnetic, +1, copper species. However, this conclusion was in conflict with the single-crystal X-ray structure, which supported the presence of copper +2. The oxidation state of the metal centre is implicitly linked to the charge of the bulk

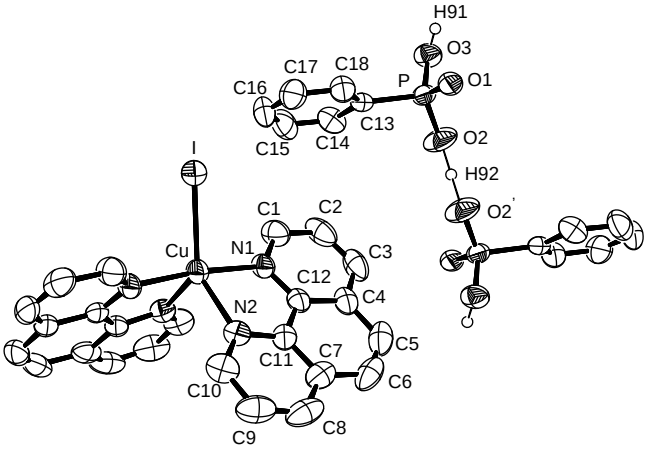

Figure 1.1: An ORTEP perspective of the molecular unit of the iodo-form, $[Cu(C_{12}H_8N_2)_2I][(OH)_2OPC_6H_5][(OH)O_2PC_6H_5]$.

structure, hence the number of protons and geometry of the hydrogen bond linking the phenylphosphonic acid molecules was also uncertain. A computational study [21] was undertaken to answer the following two questions: Are there one or two protons at the hydrogen bonding site and hence what is the oxidation state of the copper centre? Analysis of the binding energies of the one- and two-proton systems revealed that the one-proton arrangement was energetically more favourable. This arrangement is consistent with a paramagnetic Cu^{2+} centre, the presence of which was subsequently confirmed by experimental (masNMR) evidence.

Our present interest in these materials stems from the desire to better understand the complex interactions often present in hybrid organic/inorganic molecular materials. As we have already discussed, a detailed understanding of the properties of condensed molecular matter begins with an understanding of interactions *between* molecules. The diverse range of elements and bonding scenarios in the mixed-ligand organophosphonates means they provide interesting 'laboratories' for the study of intermolecular interactions.

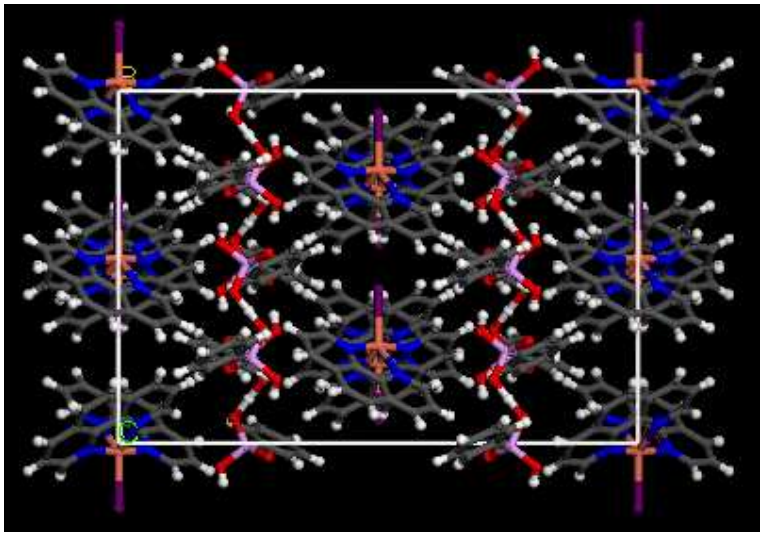

Figure 1.2: A 1×2×1 supercell of the iodo-form. Different atoms are distinguished by the following colours: grey–carbon, white–hydrogen, blue–nitrogen, orange–copper, dark purple–iodine, light purple–phosphorus, red–oxygen.

1.2 Intermolecular Interactions

The mixed-ligand organophosphonates contain a variety of interactions: hydrogen bonds between phenylphosphonic acid molecules, dispersion interactions between aromatic rings, and metal-ligand interactions between the copper ions and halogen and phenanthroline ligands. There are also electrostatic interactions between the positively charged copper complex molecules and the negatively charged phenylphosphonic acid dimers. Figures 1.3 and 1.4 are coloured according to the strength and type of interaction present. Let us briefly consider each interaction in turn.

Hydrogen Bonding

Hydrogen bonding is one of the most ubiquitous non-covalent forces leading to intricate assemblies of molecules. This interaction was first explicitly recognised by Latimer and Rodebush in 1920. [22] In those days, inferences about molecular interactions were drawn from bulk properties, such as the enthalpy of vaporisation. Major advances were made in the 1980s with Millen's gas-phase spectroscopic studies of a selection of true dimers, such as HCN$\cdots$HF. [23, 24] Theoreticians subsequently performed *ab initio* studies and a large body of work, both experimental and theoretical, has since been

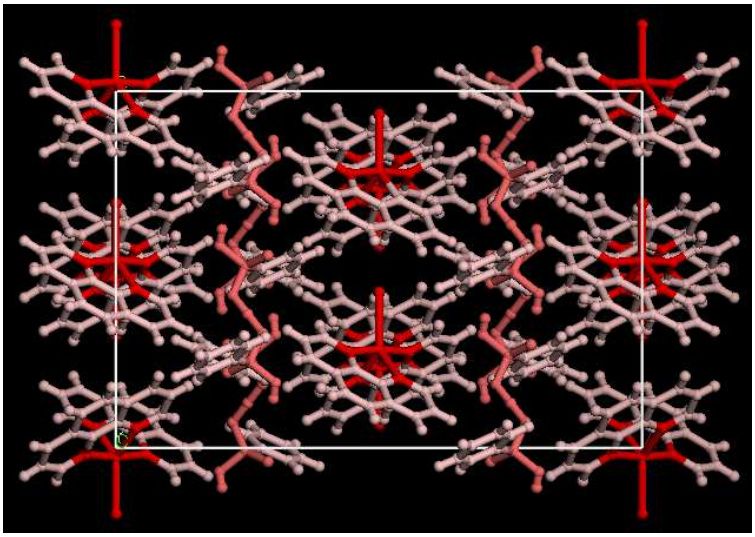

Figure 1.3: A 1×2×1 supercell coloured according to the the strength of the intermolecular interactions present. Darker shades of red indicate stronger interactions.

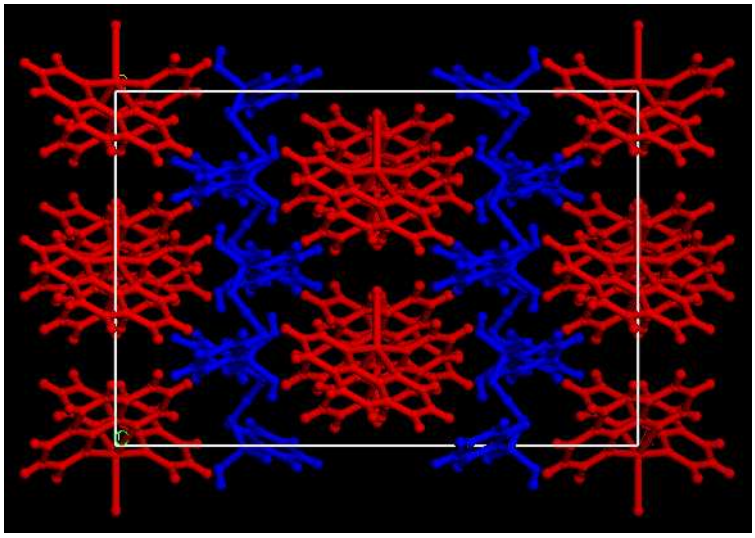

Figure 1.4: A 1×2×1 supercell showing electrostatic interactions between positively charged copper complex molecules (red) and negatively charged phenylphosphonic acid dimers (blue).

Table 1.1: Selected properties of strong, moderate and weak hydrogen bonds following the classification of Jeffrey [1].

	Strong	**Moderate**	**Weak**
Interaction type	Strongly covalent	Mostly electrostatic	Electostatic/ Dispersion
A$\cdots$B lengths, Å	2.2-2.5	2.5-3.2	> 3.2
Directionality	Strong	Moderate	Weak
Bond angles	170-180^o	$> 130^o$	$> 90^o$
Bond energy, kcal mol^{-1}	15-40	4-15	< 4

accumulated.

Hydrogen bonding interactions occur when a hydrogen atom, H, covalently bound to a heteroatom, A, is in close contact with a second electronegative atom, B, typically at a distance of 2.2–3.2 Å, *i.e.* A–H$\cdots$B. The site of attraction on B is commonly identified as a lone pair of electrons, hence hydrogen bonds are usually found to link electronegative atoms, although other, more exotic, interactions are possible: CH$\cdots$O [13], MH$\cdots$O, where M is a metal [25], MX$\cdots$H, where X is a halide [26] and even CH$\cdots\pi$, where π is an aromatic ring, such as benzene. [27] The hydrogen bond is thus a very general phenomenon with a correspondingly broad range of properties. A selection of these properties is listed in Table 1.1. The classification presented here is purely one of practical convenience. Hydrogen bonds exist in a continuum of strengths and geometries; indeed, many hydrogen bonds could be classed as 'in between' one of the categories listed above.

The theoretical description of hydrogen-bonded systems remains a challenge, partly due to the relative weakness of the interaction. A number of chemical species can participate in hydrogen bonding, which suggests that the interaction can originate from a variety of physical sources, such as electrostatic and covalent forces. An accurate theoretical description must thus include a balanced description of these subtle effects. It appears that Density Functional Theory can adequately describe the geometry and energetics of stronger hydrogen bonds but struggles for weaker bonds where the contributions from dispersion are increased. [28, 29] One certainty is that the choice of functional and basis set is important. This issue will be examined in more detail in Chapter 3.

Dispersion Interactions

Although the dispersion (or van der Waals) interaction is ubiquitous, it is very weak (< 2 kcal/mol) and is thus often masked by stronger (ionic, covalent) forces in the solid state. Typical examples of van der Waals solids include the rare-gas crystals, Ne, Ar, Kr and Xe. The interaction between two or more aromatic rings is often attributed to the dispersion force, hence organic molecular crystals such as benzene, could also be considered van der Waals solids. Dispersion forces also play an important role in many soft-matter situations, such as protein folding and polymer cohesion.

Transient moments caused by zero-point quantum fluctuations are responsible for dispersion interactions. This is in contrast to electrostatic forces, which arise when one or more of the systems possess a permanent electric moment. Consider the simplest example of the dispersion interaction, that of two well-separated atoms, A and B. Fluctuations generate a spontaneous, transient dipole moment on atom A. This dipole causes an electric field at atom B, thus inducing a dipole on B and creating a field back at A. The interaction energy between this field and the original transient dipole scales like R^{-6}, where R is the distance between A and B. The dipolar fluctuations on the two systems are correlated.

Dispersion interactions are even more difficult to describe theoretically than hydrogen bonds. The dispersion force originates from a high-order interaction and is thus not easily incorporated into theories of electronic structure. Computational schemes (such as Configuration Interaction) are available to treat the simplest case of two finite non-overlapping systems. However, these methods are considered impractical for calculations on large systems with significant dispersion energy contributions. The difficulties become increasingly severe in the case of overlapping systems, where dispersion forces compete with other types of bonding. Theoretical schemes to treat the dispersion interaction in this regime do exist (Quantum Monte Carlo methods [30] and Random Phase Approximation [31]). However, they are also computationally intensive and generally inapplicable to very large systems. The most widely-used theory of electronic structure, Density Functional Theory, is incapable of (systematically) accurately describing systems dominated by dispersion interactions, either overlapping or

well-separated. [32]

Metal-Ligand Interactions

Metal-ligand interactions are intramolecular, unlike hydrogen bonding and dispersion, which are intermolecular interactions. In the case of the organophosphonates described earlier, the relevant interactions are Cu–X and Cu–N.

M–L interactions form the basis of coordination chemistry and are usually strong and short. Bond distances are normally in the range 1.8 Å to 2.5 Å. [33] The Cu–X bond distances in the organophosphonates are in the range 1.90 Å to 2.64 Å and the Cu–N distances are in the range 1.98 Å to 2.12 Å. M–L complexes are less well characterised (both theoretically and experimentally) than their organic counterparts, hence bond energy data are scarce. Some examples (from Dance, Ref. [12]) are 40–60 kcal/mol for M–O in acetylacetonate complexes [34], 40–70 kcal/mol for Pd/Pt–N/P/Cl in amine and phosphine complexes L_2MCl_2 [35], 40–55 kcal/mol for M–S in complexes of $^{-}S_2CNEt_2$ [34], 33 kcal/mol for Ag^{+}–OH_2 [36], 52 kcal/mol for H_3NCu^{+}–NH_3 [36] and 124 kcal/mol for Zr–O in $Zr(OPr)_4$. [37]

1.3 Computational Materials Science

Crystal structures are often described in geometric terms. However, it is the energetics of intermolecular interactions that determine the chemical and physical properties of molecular materials. [12] Experimental data on the energies associated with particular interactions are scarce, mostly due to the difficulty involved in carrying out such experiments. Hence, computational simulations are an important complement to experimental techniques in probing the electronic structure of complex materials. The detailed theory of computational simulations is presented in Chapter 2. Here we briefly review the applicability, advantages and disadvantages of the two major computational techniques used in this work – Density Functional Theory and Quantum Monte Carlo.

1.3.1 Density Functional Theory

The advent of Density Functional Theory (DFT) has made possible accurate calculations on large systems containing a variety of atoms in variable oxidation and spin states. This is because DFT is, in principle, an exact theory of electronic structure and hence does not rely on empirical data. Conversely, semi-empirical and forcefield methods must be parameterised, either using experimental data or the results from high-level quantum mechanical calculations. Hence, these latter methods are generally not transferable between different types of systems. Given the enormous variety of possible interactions in molecular crystals, especially inorganic ones, an *ab initio* method such as DFT is highly desirable.

The list of systems and phenomena successfully described by DFT is impressive. With regards to molecular crystals, organic semiconductors, such as the polyacene family, have perhaps been the most extensively studied. In these studies, DFT has been used as a complement to experimental studies to describe the hole-phonon interaction in anthracene, tetracene and pentacene. [38] These results have important implications for the theoretical description of charge transport in organic semiconductors. DFT has also been used to calculate the electronic and optical properties of anthracene at high pressures. [7] The application of hydrostatic pressure provides one way to alter the intermolecular interactions in a controlled manner. The surface energetics of pentacene have been successfully described within a pseudopotential DFT approximation and the results used to formulate a model of intermolecular bonding in polyacenes. [39]

DFT has also been applied to the study of inorganic molecular materials, though there are far fewer examples in the literature. DFT calculations have been used to show that, contrary to the popular conception, short bonds are not always strong bonds. [40] Theoretical studies on iodine clusters have also been carried out. [12] Iodine readily forms complexes with amines and with itself to form a range of polyiodides. Given the size and polarisability of iodine, we could expect the intermolecular energies of I_2 to be reasonably large; this assumption has been confirmed by DFT calculations. Metal complexes with aromatic ligands are common in inorganic molecular materials yet, until the first DFT calculations were performed [12], nothing was known about

the intermolecular energies associated with these compounds; researchers in this field await experimental confirmation of the predictions made by DFT.

In addition to some stunning successes, DFT has also had its fair share of failures, including the infamous 'band gap problem' and self-interaction error. A detailed discussion of these errors (and possible solutions) is reserved for Chapter 2.

1.3.2 Quantum Monte Carlo

Powerful though DFT is, it often lacks the accuracy required to describe, for example, weak physical interactions such as van der Waals and hydrogen bonding. The Quantum Monte Carlo (QMC) method is an explicitly correlated many-body technique for solving the Schrödinger equation and it has successfully described physical phenomena in both condensed and molecular systems.

Since McMillan first used a QMC method in his study on the ground state properties of liquid helium [41], QMC has been employed in many other studies. Ceperley and Alder used QMC to calculate the correlation properties of the homogeneous electron gas [42] to a very high accuracy – simulations which subsequently lead to the development of the Local Density Approximation (LDA). A recent important advance has been the development of a "hybrid" technique in which QMC is coupled to a molecular dynamics simulation [43], *i.e.* the QMC energy is evaluated at atomic positions generated by *ab initio* molecular dynamics. The accuracy of this promising approach has been demonstrated for a variety of systems, including in calculations of the optical gaps of hot silicon quantum dots. QMC has also been used to study the electronic spectrum of free base porphyrin, an important component in numerous biological processes. [44] The excitation spectra of these systems are sensitive to structural and chemical changes hence only a highly accurate method, such as QMC, could produce reliable theoretical predictions.

Unfortunately, while QMC is highly accurate, it is also much more computationally demanding than DFT. The QMC algorithm is however trivially parallelisable and this makes QMC simulations ever more feasible as the availability of powerful supercom-

puters increases. A detailed discussion of the QMC technique is presented in Chapter 6.

1.4 Putting It All Together: Materials, Interactions, Theory

We have discussed the emergence of molecular materials as important components of new technological devices. We have also discussed the importance of understanding the interactions between the molecules that make up these materials and how this understanding can be achieved using computer simulations. Hence, the aim of this work is to use theoretical methods, such as Density Functional Theory and Quantum Monte Carlo, to probe the electronic structure of complex molecular materials in an effort to understand both the intra- and intermolecular interactions.

By way of introduction, in Chapter 2 we discuss the theoretical background behind the computational techniques used in this work. The hydrogen-bonded phenylphosphonic acid dimers under study in this work are large systems and contain heavy atoms. We therefore require an accurate and efficient computational technique. Numerical basis sets may provide the speed and accuracy required but systematic studies concerning their performance in hydrogen bonding applications are lacking. In Chapter 3, we compute the structure and binding energies of some well-characterised hydrogen-bonded systems using two different types of basis set: Gaussian-type and numerical. In the following chapters (4 and 5), we apply the chosen methodology to the study of interactions between the complex components that make up the organophosphonates. In Chapter 6, we describe and test the applicability of the Quantum Monte Carlo method to describe weakly-bound molecular systems. We take water dimer, the archetypal hydrogen bonded system, as our test case. We draw our conclusions in Chapter 7.

Chapter 2

Theory and Methods

2.1 Introduction

The theoretical methods of quantum mechanics form the basis of our understanding of the electronic structure of matter. Electronic structure calculations involve solving the Schrödinger equation for the system of interest, in order to determine the wavefunction, which contains all of the information about the given system. In practice, we are required to make several approximations to the form of the wavefunction and the Hamiltonian. In this chapter, we discuss two such approximations and thus examine the theoretical background underlying the calculations performed in this work.

2.2 The Many-Electron Problem

The many-electron problem arises when one attempts to solve the time-independent, non-relativistic Schrödinger equation

$$\hat{H}\Psi = E\Psi \tag{2.1}$$

where Ψ is the wavefunction and E is the total energy of the system. The Hamiltonian operator, $\hat{H}$, describes the motion of N electrons in the field of M fixed nuclei and has

the well known form

$$\hat{H} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{R_{iA}} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N} \frac{1}{r_{ij}} \tag{2.2}$$

in Hartree atomic units, so that $\hbar = m_e = e = 4\pi\epsilon_0 = 1$. The nuclear charge is given by Z_A, R_{iA} is the distance between electron i and nucleus A and r_{ij} is the distance between electron i and electron j.

The Born-Oppenheimer approximation [45] allows us to neglect all terms in $\hat{H}$ which arise from the repulsion and kinetic energy of the nuclei. The remaining terms are shown in (2.2) and derive from electronic motion and coulomb forces: the first term ∇_i^2 is the Laplacian operator for the kinetic energy of the electrons; the second term describes the coulomb attraction between the electrons and nuclei; and the third term represents electron-electron repulsion. It is this third term[1] which correlates the motion of the electrons and precludes an exact solution of the Schrödinger equation for systems of more than two interacting particles. This difficulty is known as the many-electron problem; the use of approximate methods to circumvent it forms the basis of electronic structure theory. The remainder of this chapter will be devoted to a discussion of two such approximate methods: Hartree-Fock theory and Density Functional Theory.

2.3 Hartree-Fock Theory

In Hartree-Fock theory [46, 47] the N-electron wavefunction is approximated by an antisymmetrised product of N one-electron wavefunctions, otherwise known as orbitals. In practice, this is achieved through the Slater determinant [48], Ψ_{HF}

$$\Psi_{\mathrm{HF}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_i(\mathbf{x_1}) & \chi_j(\mathbf{x_1}) & \cdots & \chi_k(\mathbf{x_1}) \\ \chi_i(\mathbf{x_2}) & \chi_j(\mathbf{x_2}) & \cdots & \chi_k(\mathbf{x_2}) \\ \vdots & \vdots & & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \cdots & \chi_k(\mathbf{x}_N) \end{vmatrix} \tag{2.3}$$

[1] Whilst the third term is perhaps the most physically interesting, the electron-nucleus term also prohibits exact solutions in non-trivial systems.

where the matrix elements are the spin orbitals and denote the space $\mathbf{x}$ and spin χ coordinates of the electrons. The orbitals are chosen to be orthonormal for computational convenience. Furthermore, in order to satisfy the Fermi statistics of electrons, the wavefunction is antisymmetric under exchange of any two electrons, *i.e.*

$$\Psi_{\mathrm{HF}}(\mathbf{x_1}, \mathbf{x_2}) = -\Psi_{\mathrm{HF}}(\mathbf{x_2}, \mathbf{x_1}). \tag{2.4}$$

Equation (2.4) now necessarily obeys the Pauli Principle. [49] The Slater determinant fulfills the antisymmetry condition, since a determinant changes sign upon interchange of two rows or columns.
Hence, in Hartree-Fock theory, the approximation is made right at the beginning: the true N-electron wavefunction is replaced by a single Slater determinant of N one-electron functions. While the equations that follow are exact, their forms are a direct consequence of this approximation.

In general, analytic solutions for the Schrödinger equation do not exist, even when using an approximate N-electron wavefunction. We therefore require some procedure to decide whether a particular wavefunction is a better approximation to the true solution than another. The variational principle [50, 51] can be employed to find the Slater determinant which yields the lowest single-determinant Hartree-Fock energy, E_{HF}. The variation method is based on taking a trial function $|\widetilde{\phi}\rangle$, which depends on certain parameters. The parameters are varied until the expectation value $\langle\widetilde{\phi}|\hat{H}|\widetilde{\phi}\rangle$ reaches a minimum. The variational principle also guarantees that the energy of the trial function will always be equal to or higher than the exact ground state energy. The only variational flexibility in the Slater determinant is provided by the spin orbitals. Hence, E_{HF} is minimised with respect to variations in the form of the orbitals, subject to the condition that the orbitals remain orthonormal throughout the minimisation. E_{HF} is given by,

$$E_{\mathrm{HF}} = \langle\Psi_{\mathrm{HF}}|\hat{H}|\Psi_{\mathrm{HF}}\rangle = \sum_{i}^{N}(\mathrm{i}|\hat{\mathrm{h}}|\mathrm{i}) + \frac{1}{2}\sum_{i}^{N}\sum_{j}^{N}(J_{ij} - K_{ij}) \tag{2.5}$$

where

$$(\mathrm{i}|\hat{\mathrm{h}}|\mathrm{i}) = \int \chi_{\mathrm{i}}^{*}(\mathbf{x}_1)\left\{-\frac{1}{2}\nabla^2 - \sum_{\mathrm{A}=1}^{\mathrm{M}}\frac{Z_A}{R_{1A}}\right\}\chi_{\mathrm{i}}(\mathbf{x}_1)d\mathbf{x}_1 \tag{2.6}$$

is a one-electron integral and

$$J_{ij} = \int\int |\chi_i(\mathbf{x}_1)^2|\frac{1}{r_{12}}\,|\chi_j(\mathbf{x}_2)|^2 d\mathbf{x}_1 d\mathbf{x}_2 \tag{2.7}$$

$$K_{ij} = \int\int \chi_i(\mathbf{x}_1)\chi_j^{*}(\mathbf{x}_1)\frac{1}{r_{12}}\,\chi_j(\mathbf{x}_2)\chi_i^{*}(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2 \tag{2.8}$$

are the two-electron coulomb and exchange integrals, respectively. In this instance, the term $i = j$ is allowed in (2.5).

The orthonormality of the orbitals is maintained through Lagrange's method of undetermined multipliers. Varying (2.5) with respect to χ produces a set of coupled differential equations:

$$\hat{f}(i)\chi(\mathbf{x}_i) = \varepsilon_i\chi(\mathbf{x}_i),\ i = 1, 2, \ldots, N. \tag{2.9}$$

These are the Hartree-Fock equations. They have the form of eigenvalue problems, where the Lagrange multipliers, ε_i, are the eigenvalues of the one-electron Fock operator, $\hat{f}_i$, and have the physical interpretation of orbital energies.

The Fock operator is defined as,

$$\hat{f}_i = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{R_{iA}} + V_{\mathrm{HF}}(i). \tag{2.10}$$

$V_{\mathrm{HF}}(i)$ is the Hartree-Fock potential, a one-electron operator, which replaces the complicated two-electron repulsion operator $1/r_{ij}$ in the Hamiltonian (2.2). This modification simplifies the many-electron problem and allows its approximate solution. $V_{\mathrm{HF}}(i)$ contains the following components:

$$V_{\mathrm{HF}}(\mathbf{x}_i) = \sum_{j}^{N}\left(\hat{J}_j(\mathbf{x}_1) - \hat{K}_j(\mathbf{x}_1)\right) \tag{2.11}$$

where the coulomb operator $\hat{J}$ is defined as,

$$\hat{J}_j(\mathbf{x}_1) = \int |\chi_j(\mathbf{x}_2)|^2 \frac{1}{r_{12}}\, d\mathbf{x}_2. \tag{2.12}$$

Its expectation value (the coulomb integral, (2.7)) describes the classical coulomb interaction between two interacting charge distributions.
The second term in (2.11), $\hat{K}$, is the exchange operator. Unlike the coulomb operator, $\hat{K}$ has no classical analogue and can only be defined through its effect when operating on a spin orbital:

$$\hat{K}_j(\mathbf{x}_1)\chi_j(\mathbf{x}_1) = \int \chi_j^*(\mathbf{x}_2)\frac{1}{r_{12}}\, \chi_i(\mathbf{x}_2) d\mathbf{x}_2 \chi_j(\mathbf{x}_1). \tag{2.13}$$

The expectation value of the exchange operator is given by the exchange integral, equation (2.8).

In the double summation in (2.5), the term $i = j$ is allowed. When $i = j$, the integral describes the coulomb interaction of the charge distribution of one electron with itself. Consequently, if we compute the energy of a one-electron system, (2.5) will give a non-zero result, even though there is no electron-electron repulsion. This source of error is called the unphysical self-interaction. Fortunately, the coulomb and exchange integrals have the important property,

$$J_{ii} = K_{ii} \tag{2.14}$$

so now when $i = j$, the integrals enter (2.5) with opposite signs and the self-interaction is exactly cancelled. As we shall see later (Chapter 5), this is not the case in Density Functional Theory and self-interaction error represents a major theoretical problem in Density Functional approaches.

Finally, the equations represented by (2.10) are coupled such that the Fock operator for orbital χ_i depends on all the other orbitals χ_j, that is, the Fock operator depends on its eigenfunctions. The Hartree-Fock equations are therefore non-linear and must be solved iteratively. The procedure for doing so is called the self-consistent field (SCF) method. Using some initial guess for the form of the orbitals, the Hartree-Fock equations can be

solved to yield a new set of orbitals. The resulting new set of orbitals is then used as input for the next iteration and so on until the input and output orbitals differ by less than some predetermined tolerance. The calculation is then said to have converged. In practice, the orbitals are expanded in a basis to make the problem computationally tractable. Basis sets will be discussed in more detail in the following chapter.

Assumptions and Limitations

In constructing the equation for the Hartree-Fock potential, V_{HF}, we have replaced the two-electron repulsion operator $1/r_{ij}$ in the Hamiltonian by a simpler one-electron operator. The effect of this one-electron operator is to create a mean-field potential in which the instantaneous position of a particular electron is not influenced by the presence of a neighbouring electron. As noted in the previous section, this simplification allows us to approximately solve the many-electron problem, but it is not without consequences.

Let us consider again the expression for the Hartree-Fock potential, V_{HF}

$$V_{\mathrm{HF}}(\mathbf{x}_i) = \sum_{j}^{N} \left(\hat{J}_j(\mathbf{x}_1) - \hat{K}_j(\mathbf{x}_1) \right).$$

The sum is over the orbitals, which appear explicitly in the equation for the coulomb operator, $\hat{J}$

$$\hat{J}_j(\mathbf{x}_1) = \int |\chi_j(\mathbf{x}_2)|^2 \frac{1}{r_{12}} \, d\mathbf{x}_2.$$

The above equation has the following physical interpretation: orbital j containing electron 1 is operated on by $\hat{J}$, which contains an integration over the remaining electrons in their respective orbitals, thus producing the mean-field. It does not matter which of the electrons – 'number' 1 or 'number' 2 – is associated with the orbitals. The numerical labelling of the electrons is merely an artefact of the notation, since electrons are indistinguishable and therefore cannot be labelled. Conversely, the orbitals *are* distinguishable, so each orbital is held 'fixed' with respect to a reference electron while the spatial positions of each remaining electron are summed over. The reference electron then feels the average potential produced by the remaining electrons.

This mean-field approach affects interactions between antiparallel spin electrons. Electrons of parallel spin are kept apart as a consequence of the Pauli Principle and this is taken into account in the construction of the Slater determinant. This effect is known as *exchange* or *Fermi correlation* and is directly connected to the spin of the electrons; the charge of the electrons plays no part in Fermi correlation. In contrast, antiparallel spin electrons repel each other through the coulomb force and this effect (known as *coulomb correlation*, or simply *correlation*) *is* a consequence of the electronic charge. In Hartree-Fock theory, antiparallel spin electrons move independently of each other and are allowed to approach too closely. As a result, the electron-electron repulsion term is too large and the Hartree-Fock energy is usually far above the exact ground-state energy.

Despite the seemingly drastic approximation made to the wavefunction (and consequently, the coulomb interaction), Hartree-Fock theory captures much of the essential physics required to describe electron interactions and gives sensible results for many-body problems. However, Hartree-Fock models are usually only qualitatively correct and cannot accurately predict the results of experiments. Higher level methods are available, such as Moller-Plesset perturbation theory, Configuration Interaction (CI) and Coupled Cluster (CC). However, perturbation methods can suffer from convergence problems and both CI and CC are prohibitively expensive for large systems. An alternative approach, which is both accurate and economical, is therefore needed; this is provided by Density Functional Theory.

2.4 Density Functional Theory

In Density Functional Theory (DFT), the single-particle electron density $\rho(\mathbf{r})$ replaces the complicated N-electron wavefunction as the central object. As a prelude to a full discussion of the theory, we should first consider whether it seems plausible to replace the wavefunction by the electron density.

In the probability interpretation of quantum theory, the density is related to the wave-

function by the following integral,

$$\rho(\mathbf{r}_1) = N \int \cdots \int |\Psi(\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_N})|^2 ds_1 d\mathbf{x_2} \ldots d\mathbf{x_N}. \tag{2.15}$$

Equation (2.15) is really a probability density but it is commonly referred to as the electron density. The electron density then is a non-negative function of three spatial variables and has the following properties:

$$\rho(\mathbf{r} \to \infty) = 0, \tag{2.16}$$

$$\int \rho(\mathbf{r}) d\mathbf{r} = N. \tag{2.17}$$

Another important property of the electron density is that $\rho(\mathbf{r})$ exhibits cusps at the positions of the nuclei. The properties of these cusps are related to the nuclear charge Z through

$$\lim_{R_{iA} \to 0} \left[\frac{\partial}{\partial r} + 2Z_A \right] \overline{\rho}(\mathbf{r}) = 0 \tag{2.18}$$

where $\overline{\rho}(\mathbf{r})$ is the spherical average of $\rho(\mathbf{r})$.
In summary, the electron density is a function of only three spatial variables (2.15), vanishes at infinity (2.16) and integrates to the total number of electrons (2.17). It exhibits cusps at the positions of the nuclei and the density at these cusps contains information about the nuclear charge (2.18). Hence, it does indeed seem plausible that we could, at least in principle, use the electron density to obtain information about the properties of atoms, molecules and condensed matter.

Thomas [52] and Fermi [53] made the first attempts to obtain information about the ground state of atoms using the electron density. Their approach is based on a quantum statistical model of electrons in which nuclear-electron and electron-electron interactions are treated classically and the expression for the kinetic energy is based on the uniform electron gas. These attempts ultimately failed, mostly because the kinetic energy expression employed in the model is a poor approximation to the true kinetic energy. However, Thomas and Fermi were the first to show that the energy of an atom could be written entirely in terms of the electron density, without recourse to the

wavefunction.

An accurate, workable prescription for DFT was provided by the pioneering work of Kohn and collaborators in the 1960's. All modern density functional theories are based on the theorems proven by Kohn, Hohenberg and Sham:

- The First Hohenberg-Kohn Theorem [54];
- The Second Hohenberg-Kohn Theorem [54];
- The Kohn-Sham Ansatz [55].

2.4.1 The First Hohenberg-Kohn Theorem

The First Hohenberg-Kohn Theorem states that the external potential $V_{\text{ext}}(\mathbf{r})$ uniquely determines the ground state density. Since $V_{\text{ext}}(\mathbf{r})$ fixes the Hamiltonian, the full, many-particle ground state is a unique functional of $\rho(\mathbf{r})$. The proof of this theorem by *reductivo ad absurdum* is surprisingly simple.

We start from the Hamiltonian for N non-relativistic, interacting electrons in a non-magnetic state,

$$\hat{H} = \hat{T} + \hat{V}_{\text{ext}} + \hat{V}_{\text{ee}} \tag{2.19}$$

where

$$\hat{T} = -\frac{1}{2}\sum_j \nabla_j^2, \quad \hat{V}_{\text{ext}} = \sum_j v(r_j), \quad \hat{V}_{\text{ee}} = \frac{1}{2}\sum_{i\neq j}\frac{1}{|r_i - r_j|}. \tag{2.20}$$

The electrons move in an arbitrary external potential $v(r)$, not only the coulomb potential due to the nuclei.[2]

The energy can thus be written as,

$$E = \int V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + \langle\Psi|\hat{T} + \hat{V}_{\text{ee}}|\Psi\rangle \tag{2.21}$$

where $\rho(\mathbf{r})$ is the non-degenerate[3] ground state density of Ψ corresponding to the

[2] We consider only those densities which are associated with an antisymmetric wavefunction due to a Hamiltonian of the form (2.19) with *some* external potential V. These densities are said to be V-representable.

[3] The proof can be readily extended to degenerate cases.

potential $V_{\text{ext}}(\mathbf{r})$.

Now assume that there exists a second potential $V'_{\text{ext}}(\mathbf{r})$ with ground state Ψ', which differs from $V_{\text{ext}}(\mathbf{r})$ by more than an additive constant but gives rise to the same $\rho(\mathbf{r})$. That is, if Ψ' is the ground state wavefunction corresponding to $V'_{\text{ext}}(\mathbf{r})$, then the energy of this state is given by,

$$E' = \int V'_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + \langle\Psi'|\hat{T} + \hat{V}_{\text{ee}}|\Psi'\rangle. \tag{2.22}$$

Using the variational theorem, we find that

$$\begin{aligned} E &< \langle\Psi'|\hat{H}|\Psi'\rangle \\ &= \int V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + \langle\Psi'|\hat{T} + \hat{V}_{\text{ee}}|\Psi'\rangle \\ &= E' + \int [V_{\text{ext}}(\mathbf{r}) - V'_{\text{ext}}(\mathbf{r})]\rho(\mathbf{r})d\mathbf{r}. \end{aligned} \tag{2.23}$$

Similarly for E', we find that

$$E' \leqslant \langle\Psi|\hat{H}|\Psi\rangle = E + \int [V'_{\text{ext}}(\mathbf{r}) - V_{\text{ext}}(\mathbf{r})]\rho(\mathbf{r})dr. \tag{2.24}$$

Adding equations 2.23 and 2.24 leads to

$$E + E' < E + E', \tag{2.25}$$

which is clearly a contradiction. Hence, the assumption that there exists a second potential $V'_{\text{ext}}(\mathbf{r})$, which differs from $V_{\text{ext}}(\mathbf{r})$ but produces the same density must be incorrect.

2.4.2 The Second Hohenberg-Kohn Theorem

The Second Hohenberg-Kohn Theorem is a re-statement of the variational principle in terms of trial densities $\tilde{\rho}(\mathbf{r})$ rather than trial wavefunctions, $\tilde{\Psi}$. We will discuss an alternative, broader[4] approach – the constrained search method – due to Levy [56, 57] and Lieb. [58, 59]

[4]The constrained search technique lifts the requirement of V-representability, for example.

Recall the variational principle, introduced in the previous section on Hartree-Fock theory,

$$E = \min_{\tilde{\Psi}} \langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle. \tag{2.26}$$

where $\tilde{\Psi}$ is a trial function. The minimisation is carried out in two stages. First, we consider the energy for the subset of antisymmetric wavefunctions $\tilde{\Psi}^{\alpha}_{\tilde{\rho}(\mathbf{r})}$ that have the same density, $\tilde{\rho}(\mathbf{r})$. The constrained energy minimum is defined as,

$$\begin{aligned} E_V[\tilde{\rho}(\mathbf{r})] &= \min_{\alpha} \langle \tilde{\Psi}^{\alpha}_{\tilde{\rho}(\mathbf{r})} | \hat{H} | \tilde{\Psi}^{\alpha}_{\tilde{\rho}(\mathbf{r})} \rangle \\ &= \int V_{\text{ext}} \tilde{\rho}(\mathbf{r}) d\mathbf{r} + F[\tilde{\rho}(\mathbf{r})] \end{aligned} \tag{2.27}$$

where

$$F[\tilde{\rho}(\mathbf{r})] = \min_{\alpha} \langle \tilde{\Psi}^{\alpha}_{\tilde{\rho}(\mathbf{r})} | \hat{T} + \hat{V}_{\text{ee}} | \tilde{\Psi}^{\alpha}_{\tilde{\rho}(\mathbf{r})} \rangle \tag{2.28}$$

The expression defined in (2.28) is the Hohenberg-Kohn functional, a universal[5] functional of the density $\tilde{\rho}(\mathbf{r})$, which requires no explicit knowledge of $V_{\text{ext}}(\mathbf{r})$. The functionals for the kinetic energy and for the electron-electron interaction are contained within (2.28). If the explicit form for $F[\tilde{\rho}(\mathbf{r})]$ were known, we would have solved the Schrödinger equation exactly. Unfortunately, the form of $F[\tilde{\rho}(\mathbf{r})]$ is not known and we are forced to make some approximations.

In the second step of the constrained search method, (2.27) is minimised over all possible densities, $\tilde{\rho}(\mathbf{r})$:

$$\begin{aligned} E &= \min_{\tilde{\rho}(\mathbf{r})} E_V[\tilde{\rho}(\mathbf{r})] \\ &= \min_{\tilde{\rho}(\mathbf{r})} \left\{ \int V_{\text{ext}} \tilde{\rho}(\mathbf{r}) d\mathbf{r} + F[\tilde{\rho}(\mathbf{r})] \right\}. \end{aligned} \tag{2.29}$$

We have thus reduced the formidable problem of minimising $\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle$ with respect to the $3N$-dimensional trial function $\tilde{\Psi}$ to the much simpler task of finding the minimum of $E_V[\tilde{\rho}(\mathbf{r})]$ with respect to the three-dimensional trial function $\tilde{\rho}(\mathbf{r})$.

[5] In this sense, "universal" means the same for all electron systems, independent of N, R_A and Z_A.

Assumptions and Limitations

Before discussing the Kohn-Sham approach, we should pause to summarise and consider some of the consequences of the First and Second Hohenberg-Kohn Theorems.

The First Hohenberg-Kohn Theorem proves that the external potential uniquely determines the density and the Second Hohenberg-Kohn Theorem (through the constrained search method) provides a route from the ground state density to the ground state energy. Unfortunately, neither theorem provides any information as to how the functional that delivers the ground state energy should be constructed. Furthermore, the variational principle established by the Second Hohenberg-Kohn Theorem is not really a variational principle in the same sense as Hartree-Fock. Conventional wavefunction methods employ the exact electronic Hamiltonian and are strictly variational: the lower the energy, the closer the trial wavefunction is to the true wavefunction. In DFT however, we are forced to use an approximation to the universal functional, which is analogous to using an approximate Hamiltonian. Hence, the DFT energy can be lower or higher than the true ground state energy.

The Second Hohenberg-Kohn theorem – as realised in the constrained search approach – also suffers from the fact that the scheme is impossible to implement in practice. The first step in the constrained search technique requires a search over all antisymmetric wavefunctions corresponding to a particular trial density (2.27). Carrying out such a search is simply not possible. It follows that the Hohenberg-Kohn functional $F[\tilde{\rho}(\mathbf{r})]$ is thus rendered inaccessible, since it explicitly relies on the wavefunction – precisely the object we wish to avoid.

2.4.3 The Kohn-Sham Ansatz

In the Thomas-Fermi approach, much of the error lies in the approximation of the kinetic energy. In the previous sections, we saw that the kinetic energy functional in DFT is defined through $F[\tilde{\rho}(\mathbf{r})]$, the universal but unknown functional of the density. The challenge then is to formulate an accurate expression for the kinetic energy of a fully interacting many-body system with the electron density as the fundamental

quantity.

Kohn and Sham realised that if the kinetic energy couldn't be accurately determined through an explicit functional, perhaps a better approach would be to compute as much of the kinetic energy exactly as possible. To accomplish this, they replaced the fully interacting system with a non-interacting reference system with the same density as the original. The reference system consists of N non-interacting electrons, that is, uncharged fermions. In formal terms, the Kohn-Sham Ansatz assumes that the exact ground state density is equal to that of some chosen non-interacting system. The practical consequence of this assumption is a set of single-particle equations that can be solved exactly (numerically in practice, like in HF theory) to arrive at the ground state energy.

The Hamiltonian corresponding to the non-interacting system is given by,

$$\hat{H}_S = -\frac{1}{2}\sum_i^N \nabla_i^2 + \sum_i^N V_{\mathrm{S}}(\mathbf{r}) \tag{2.30}$$

where we have introduced an effective local potential $V_{\mathrm{S}}(\mathbf{r})$, in which the S subscript denotes that we are working with the non-interacting reference system. The Slater determinant is constructed in exactly the same way as in Hartree-Fock theory (equation 2.3). The spin orbitals are determined by,

$$\hat{f}^{KS}\varphi_i = \varepsilon_i\varphi_i, \tag{2.31}$$

where φ_i are the Kohn-Sham orbitals and the one-electron Kohn-Sham operator is defined as,

$$\hat{f}^{KS} = -\frac{1}{2}\nabla^2 + V_{\mathrm{S}}(\mathbf{r}). \tag{2.32}$$

Equation (2.31) has the same form as the Hartree-Fock equations, (2.9). For an appropriate choice of $V_{\mathrm{S}}(\mathbf{r})$, the non-interacting system has the required properties.

Now that we have introduced orbitals, the kinetic energy can be computed in the same

way as Hartree-Fock theory

$$T_{\mathrm{S}} = -\frac{1}{2}\sum_{i}^{N}\int \varphi_i^* \nabla^2 \varphi_i, \tag{2.33}$$

where T_{S} is the kinetic energy of the non-interacting system. Although T_{S} is uniquely defined for any density, it does not exactly equal the kinetic energy of the fully interacting system. We can account for this by separating the functional $F[\rho(\mathbf{r})]$ into the following components:

$$F[\rho(\mathbf{r})] = T_{\mathrm{S}}[\rho(\mathbf{r})] + J[\rho(\mathbf{r})] + E_{\mathrm{xc}}[\rho(\mathbf{r})] \tag{2.34}$$

where E_{xc} is the exchange-correlation functional and is defined through (2.34) as

$$E_{\mathrm{xc}}[\rho] = T[\rho(\mathbf{r})] - T_{\mathrm{S}}[\rho(\mathbf{r})] + V_{ee}[\rho(\mathbf{r})] - J[\rho(\mathbf{r})] \tag{2.35}$$

The exchange-correlation functional thus contains the difference between the interacting and non-interacting kinetic energies and V_{ee}, the non-classical part of the electron-electron interaction. If the exact form for E_{xc} were known, the Kohn-Sham approach would lead to the exact energy, that is to say, the formalism discussed so far does not involve any approximations. This is in contrast to Hartree-Fock theory, where the approximation is introduced at the very beginning. Unfortunately, the exact form for E_{xc} is unknown so we are forced to approximate it. Formulating accurate approximations to E_{xc} is a major challenge for density functional theorists.

It was earlier mentioned that the link between the non-interacting and fully interacting systems lay in the choice of $V_{\mathrm{S}}(\mathbf{r})$, the effective local potential. We need to define $V_{\mathrm{S}}(\mathbf{r})$ so that the density of the non-interacting reference system is exactly the same as the system of interest. Since T_{S} is an explicit functional of the orbitals but all other terms are functionals of the density, we can write down an equation for φ_i using the

variational principle, in the same manner as for Hartree-Fock,

$$\left(-\frac{1}{2}\nabla^2 + \left[\int \frac{\rho(\mathbf{r}_2)}{r_{12}}\, d\mathbf{r}_2 + V_{\mathrm{xc}}(\mathbf{r}_1) - \sum_A^M \frac{Z_A}{r_{1A}}\right]\right)\varphi_i$$
$$= \left(-\frac{1}{2}\nabla^2 + V_{\mathrm{eff}}(\mathbf{r}_1)\right)\varphi_i = \varepsilon_i\varphi_i \tag{2.36}$$

where V_{xc} is the exchange-correlation potential. The orbitals are subject to the usual orthonormality constraint. Recalling the one-particle equations of the non-interacting system, (2.32), it becomes clear that $V_{\mathrm{eff}}(\mathbf{r})$ is identical to $V_{\mathrm{S}}(\mathbf{r})$ and we can write,

$$V_{\mathrm{S}}(\mathbf{r}) = V_{\mathrm{eff}}(\mathbf{r}) = \int \frac{\rho(\mathbf{r}_2)}{r_{12}}\, d\mathbf{r}_2 + V_{\mathrm{xc}}(\mathbf{r}_1) - \sum_{A=1}^{M} \frac{Z_A}{R_{1A}}. \tag{2.37}$$

Since the form of E_{xc} is unknown, it follows that the expression for V_{xc} is also unknown. Hence, V_{xc} is given by the functional derivative of E_{xc} with respect to the density,

$$V_{\mathrm{xc}} = \frac{\delta E_{\mathrm{xc}}}{\delta\rho}. \tag{2.38}$$

In summary, we first define a non-interacting reference system, the exact ground state energy of which is a single Slater determinant and the density of which exactly equals the density of the fully interacting system. The orbitals forming the Slater determinant are the solutions of a set of single-particle equations (2.31). Once the orbitals have been determined, the kinetic energy of the non-interacting system can be computed in the same way as for Hartree-Fock theory (2.33). The effective potential, which must be chosen to give the correct density for the non-interacting system, can be calculated self-consistently by solving the set of variational equations (2.36). The orbitals, the density and hence the ground state energy can thus be found. If the explicit forms of all the potentials were known, the orbitals would return the exact density and energy. In practice, we must resort to approximations for the exchange-correlation functional, E_{xc}.

2.4.4 Functionals for Exchange and Correlation

The crucial quantity in the Kohn-Sham approach is the exchange-correlation energy, expressed as a functional of the density $E_{\mathrm{xc}}[\rho(\mathbf{r})]$. The difference between various DFT methods essentially lie in the choice of the functional form for $E_{\mathrm{xc}}[\rho(\mathbf{r})]$.

At the heart of almost all current functionals is the Local Spin Density Approximation (LSDA),[6]

$$E_{\mathrm{xc}}^{\mathrm{LSDA}}[\rho^{\uparrow}(\mathbf{r}), \rho^{\downarrow}(\mathbf{r})] = \int \rho(\mathbf{r}) e_{\mathrm{xc}}(\rho^{\uparrow}(\mathbf{r}), \rho^{\downarrow}(\mathbf{r}))\, d\mathbf{r} \tag{2.39}$$

where e_{xc} is the the exchange-correlation energy per particle of a homogeneous electron gas (HEG) of density $\rho(\mathbf{r})$. The HEG is a hypothetical model which describes a system of electrons moving on a positively charged background such that the entire system is electrically neutral. The quantity e_{xc} can be further split into an exchange part and a correlation one,

$$e_{\mathrm{xc}}(\rho^{\uparrow}(\mathbf{r}), \rho^{\downarrow}(\mathbf{r})) = e_{\mathrm{x}}(\rho^{\uparrow}(\mathbf{r}), \rho^{\downarrow}(\mathbf{r})) + e_{\mathrm{c}}(\rho^{\uparrow}(\mathbf{r}), \rho^{\downarrow}(\mathbf{r})). \tag{2.40}$$

A simple analytic form is available for the exchange term, $e_{\mathrm{x}}\rho(\mathbf{r})$, however no such expression exists for the correlation part, $e_{\mathrm{c}}\rho(\mathbf{r})$. Ceperley and Alder [42] have calculated $e_{\mathrm{c}}\rho(\mathbf{r})$ to very high accuracy ($\pm 1\%$) using Quantum Monte Carlo methods. These results have formed the basis of various analytical expressions for $e_{\mathrm{c}}\rho(\mathbf{r})$, *i.e.* correlation functionals.

Considering the nature of the HEG, we would thus expect the LSDA to perform well for systems in which the charge density is uniform, such as nearly-free-electron metals, and poorly for strongly inhomogeneous systems, such as atoms and molecules. Despite these expectations, the LSDA performs remarkably well for both solids and molecules. The dissociation energies of molecules can typically be obtained to within 10-20%. However, most chemical applications (which frequently involve weak molecular bonds) require greater accuracy than the LSDA can provide.

The first step beyond the LSDA involves taking into account not only the density

[6] For unpolarised systems, the LDA is found by setting $\rho^{\uparrow}(\mathbf{r}) = \rho^{\downarrow}(\mathbf{r}) = \rho(\mathbf{r})/2$.

$\rho(\mathbf{r})$ at each point, but also the gradient of the charge density, $\nabla\rho(\mathbf{r})$. Within this approximation, E_{xc} is given by,[7]

$$\begin{aligned} E_{\mathrm{xc}}^{\mathrm{GGA}}[\rho^{\uparrow},\rho^{\downarrow}] &= \int e_{\mathrm{xc}}(\rho^{\uparrow},\rho^{\downarrow},|\nabla\rho^{\uparrow}|,|\nabla\rho^{\downarrow}|,\ldots)\rho\, d\mathbf{r} \\ &\equiv \int e_{\mathrm{x}}(\rho)F_{\mathrm{xc}}(\rho^{\uparrow},\rho^{\downarrow},|\nabla\rho^{\uparrow}|,|\nabla\rho^{\downarrow}|,\ldots)\rho\, d\mathbf{r} \end{aligned} \tag{2.41}$$

where F_{xc} is a dimensionless parameterised analytic function (different GGA functionals employ different forms for F_{xc}) and e_{x} is the exchange energy of the unpolarised HEG. Because only first derivatives of the density are employed, this model can be used to describe systems where the density is slowly varying. Functionals constructed with this modification are collectively known as Generalised Gradient Approximations (GGA). GGAs yield improved atomisation and dissociation energies and a more accurate description of molecular systems, particularly those in which weak interactions (hydrogen bonds) are present. GGAs have also been shown to yield improved (compared to LSDA) cohesive energies for solids.

There are now many GGA functionals that are used in quantitative calculations. The PBE functional used in Chapters 4 and 5 of this work was designed to be widely applicable for both quantum chemistry and condensed matter physics. The exchange part is similar to the Becke formula [60] while the correlation part is closely related to the mathematically complex PW91 functional. PBE is numerically reliable and is thus often used in DFT calculations. The form of the PBE functional is quite complicated, hence we refer the interested reader to the original reference. [61]

There are still other classes of functional, such as the "hybrids" and orbital-dependent functionals. The latter are particularly useful in the study of strongly correlated electron systems, such as transition metal oxides and high T_c superconductors. Since neither of these types of functional are used in this work, we won't consider them further.

[7] For the sake of clarity, here we denote the density $\rho(\mathbf{r})$ as simply ρ.

Assumptions and Limitations

The most serious limitation of DFT is the need to approximate the unknown exchange-correlation functional, E_{XC}: the better the approximation, the more closely we approach the exact energy. Unfortunately, there is no systematic strategy for improving the approximation and the design of approximate functionals often involves a large degree of 'trial and error'. Despite these serious complications, much progress has been made in the field of approximate DFT, although difficulties remain.

Among the most notable failures of DFT is the infamous 'band gap problem' whereby the one-electron energy gaps of insulators and semiconductors are systematically predicted to be too small, sometimes by as much as 100%. This is caused by a failure to account for a discontinuity in the exact exchange-correlation potential upon addition or removal of an electron [62] and is inherent in ground-state DFT band structure, *i.e.* the true exchange-correlation potential must change discontinuously and this is not the case, even in the exact Kohn-Sham theory. Finding solutions to the band gap problem is currently an active area of DFT research. One promising approach involoves the use of orbital-dependent functionals, in which the required properties occur automatically.

Another serious problem in DFT is the self-interaction error, which subsequently complicates the dissociation of odd-electron bonds. Recall the expression for the exchange-correlation functional, (2.35). The classical electrostatic repulsion term, $J[\rho(\mathbf{r})]$, contains the interaction of the density with itself. In HF theory, the self-interaction is exactly cancelled by the exchange term (see Section 2.3). In DFT however, approximations to the exchange-correlation energy are made independent of $J[\rho(\mathbf{r})]$ and thus exact cancellation is not possible. It follows that in DFT, the electrostatic term will not vanish, even for a one-electron system. None of the currently available exchange-correlation functionals is free of self-interaction error. In an attempt to solve this problem, Perdew and Zunger have proposed a self-interaction corrected (SIC) form of approximate functional. [63] The reader is referred to the original reference for details of this approach. Here we merely note that the implementation of the Perdew-Zunger scheme is highly non-trivial and has therefore been of limited practical use.

Chapter 3

Testing Numerical Basis Sets for Hydrogen Bonding Applications

3.1 Introduction

Whilst the choice of functional plays a crucial part in the accuracy of a given result, the choice of basis set used to represent the orbitals is just as important. [64] The issue of basis set dependence within DFT is frequently overlooked by many researchers. Such an omission can lead to very serious errors, especially when a basis set is used in conjunction with a method for which it was not optimised. For example, two of the most popular types of basis set, the Pople sets [65] and Dunning's correlation consistent basis sets [66], were not developed for DFT yet are routinely employed in such calculations. The Pople sets were originally optimised with the Hartree-Fock method and Dunning's sets were developed for wavefunction-based *ab initio* methods.

In this chapter we have investigated and compared the ability of numerical and Gaussian-type basis sets to accurately describe the geometries and binding energies of a selection of hydrogen bonded systems that are well studied, both theoretically and experimentally. The structures under study in this thesis contain molecules linked by extensive hydrogen bonded networks and we wish to study clusters of varying size. We

therefore require an accurate, efficient method to describe large systems with heavy atoms. The numerical basis sets implemented in the DFT package DMol^3 may provide the speed and accuracy required but systematic studies concerning the use of numerical basis sets in hydrogen bonding applications are, to the best of our knowledge, non-existent.

3.2 The Basis Set Game

The Kohn-Sham equations (2.31) represent a complicated system of coupled integro-differential equations. We therefore require a computationally efficient approach to solving these equations. This is usually achieved through a linear variational procedure in which the orbitals are expanded in a set of q pre-defined basis functions $\{\eta_q\}$:

$$\varphi_i = \sum_{i=1}^{q} c_{iq}\eta_q. \tag{3.1}$$

A complete set $\{\eta_q\}$ would correspond to an infinite number of basis functions, $q = \infty$. In practice, we use a truncated basis set, which introduces the basis set superposition and incompleteness errors. Equation (3.1) can now be solved self-consistently in a similar manner to the Kohn-Sham (and Hartree-Fock) equations.

The procedure represented by (3.1) is often inaccurately described by the phrase Linear Combination of Atomic Orbitals (LCAO). Strictly speaking, 'atomic orbitals' are solutions of the Hartree-Fock equations for the atom, *i.e.* a wavefunction for a single atomic electron in a self-consistent field. Basis functions are thus *not* atomic orbitals. As we shall see below, many different types of function can be used in the expansion, with properties chosen for computational convenience.

Most commonly available *ab initio* simulation packages employ Gaussian functions (molecular applications) or plane waves [67] (condensed matter) for the expansion represented by (3.1). In addition to these popular types of basis set, others also exist: wavelets [68], Slater-type orbitals (STO's) [69] Wannier functions [70] and numerical basis sets [71, 72], although these have enjoyed considerably less attention. Numerical

basis sets as implemented in the DFT package DMol^3 [71] offer significant advantages over the more popular types of basis set. Whilst DMol^3 has been successfully employed in the study of various phenomena [73, 74, 75, 76, 12], systematic studies concerning the performance of numerical basis sets within DFT are scarce. Investigations concerning the performance of GTO's in hydrogen-bonding applications abound ([77, 78, 79, 80, 64, 81], to list but a few studies) but to the best of our knowledge, there have been only two such studies with numerical basis sets: one which evaluated the performance of numerical basis sets in DFT calculations on sulfur-containing molecules [82] and another which evaluated the performance of DMol^3 with regards to electron affinities. [73] Hence, nothing is known about the performance of these numerical bases in relation to hydrogen bonding.

3.3 Computational Details

3.3.1 Theory of Numerical Basis Sets

As with GTO's, numerical basis functions are expressed as the product of a radial function and a spherical harmonic. The radial part is obtained by numerically solving the atomic DFT equations, which along with the $-\frac{\nabla^2}{2}$ terms (required for the evaluation of the kinetic energy), are represented as a set of cubic spline co-efficients, thereby making them piecewise analytic. This representation makes the generation of analytic energy gradients much easier. Furthermore, Delley claims [71] that molecules can be dissociated exactly into their constituent atoms (within the DFT context) thereby minimising or even eliminating basis set superposition error (BSSE) (indeed no correction procedure is implemented in the DMol^3 package). Hence, it should be possible to obtain an accurate description, even for weak bonds. These sets are computationally efficient, a feature which arises from the confinement of the LDA atomic basis sets (from which the numerical basis set is constructed) within a cut-off value, r_c. [83] Strict localisation of the basis set within r_c is ensured via a so-called soft-confinement potential. The derivatives at r_c are continuous. The basis sets implemented in DMol^3 were constructed specifically for use in DFT calculations.

3.3.2 System, Functional and Basis Set Selection

The following complexes have been selected for this study: $(HF)_2$, $(H_2O)_2$, $(HCl)_2$, (CO)HF, (OC)HF and phosphinic acid dimer, $[H_2PO(OH)]_2$. They have been chosen because of the wealth of experimental and theoretical data available. Although one could consider a larger set of systems, we believe the current set is suitably diverse so as to demonstrate the behaviour of the numerical basis sets in a range of different bonding situations.

A plethora of functionals now exist in DFT but only 'pure' functionals are implemented in the DMol^3 package, *i.e.* no hybrid functionals, such as B3LYP, are available. The functionals chosen were BLYP, [60, 84, 85] PBE [61] and HCTH fit to 407 molecules. [86] Both BLYP and HCTH are semi-empirical functionals obtained by fitting parameters to a so-called"training set" of molecules. Gaussian03 [87] was used for all calculations with Gaussian basis sets.

Both the Pople sets and Dunning sets were used as well as the numerical sets. The Gaussian sets are in widespread use amongst the electronic structure community and have been implemented in a number of quantum chemistry codes. Dunning's correlation-consistent triple-zeta quality set augmented with diffuse functions was used. This set is similar in quality to the TZ2P set [88], which has been shown to give good results within DFT for the G2-1 and G3 test sets. [64] Pople's 6-311+G(3df,2pd) set was used, which has also been shown to give adequate results in DFT. The DNP numerical basis set was used as it is the highest quality set available in DMol^3. This set contains a polarization d-function on heavy atoms and a polarisation p-function on hydrogen. It is comparable in size to the 6-31G** basis set however, numerical basis sets of a given size are much more accurate than Gaussian basis sets of the same size. [83] Calculations were also performed with Dunning's double-zeta basis set, the 6-311+G(2d,2p) set and the DND and DN numerical basis sets although, for the sake of conciseness, only the results for the largest basis sets have been presented here.

All geometries were optimised with each functional and basis set using Direct Inversion of the Iterative Subspace (DIIS) [89] to aid convergence. Binding energies and

vibrational frequencies were calculated as in Ref. [28]. The binding energies and bond lengths have been compared to both experimental values (where available) and the best *ab initio* result available (usually Coupled Cluster theory results). Where the best *ab initio* energy has been calculated using a fixed geometry from a previous calculation, the energy associated with that geometry from the initial calculation has also been cited, *i.e.* a single calculation does not always simultaneously produce accurate energies and accurate geometries.

Counterpoise corrections [90] have been calculated for the Gaussian basis sets. We have chosen to include both corrected and uncorrected energies in order to bring about a more consistent comparison.

It is important to be realistic about the level of accuracy attainable with DFT. Many electronic structure theorists (including myself) assess the quality of their DFT calculations by comparing the results against high-level methods, such as Coupled Cluster. While DFT performs adequately for energies and geometries it cannot, at present, systematically approach the accuracy offered by techniques such as Coupled Cluster, the W*n* family of methods or Quantum Monte Carlo (QMC). Hence, we have not investigated whether or not the numerical basis sets can reproduce the complete basis set limit on the grounds that a DFT result extrapolated to the complete basis set limit is unlikely to yield a more accurate result than a Coupled Cluster or QMC calculation. While it may be important, in a general sense, for a given type of basis set to be able to reproduce this limit, we have adopted a pragmatic approach and have only considered the behaviour of the numerical basis sets in the limit in which we are interested, *i.e.* geometry optimisation of large systems.

3.4 Results and Discussion

Table 3.1 shows bond lengths and vibrational stretching frequencies for the isolated proton donating subunits for HF, HCl and H_2O monomers. A general trend is apparent for all three systems.

Bond lengths are overestimated in all cases, regardless of the basis set. This is a char-

Table 3.1: Equilibrium geometries and stretching frequencies for isolated proton donating subunits. All bond lengths in Å, frequencies in cm^{-1}.

Functional/Basis	**HF**		**HCl**		**H_2O**			
	r_{H-F}	ω	r_{H-Cl}	ω	r_{O-H}	ω_{as}	ω_s	ω_f
BLYP								
DNP	0.9400	3951	1.2920	2873	0.9730	3791	3677	3735
6-311+G(3df,2pd)	0.9408	3938	1.2902	2847	0.9700	3785	3687	3736
aug-cc-pVTZ	0.9350	3913	1.2934	2840	0.9720	3757	3655	3706
HCTH								
DNP	0.9230	4078	1.2760	3028	0.9620	3927	3804	3866
6-311+G(3df,2pd)	0.9193	4088	1.2746	2971	0.9590	3927	3816	3872
aug-cc-pVTZ	0.9217	4062	1.2799	2956	0.9600	3897	3785	3841
PBE								
DNP	0.9330	4008	1.2900	2899	0.9710	3849	3731	3791
6-311+G(3df,2pd)	0.9300	3986	1.2878	2896	0.9680	3829	3726	3777
aug-cc-pVTZ	0.9321	3962	1.2911	2886	0.9700	3801	3697	3749
Best *ab initio*	0.9173[1]	4142[1]	1.2783[2]	2999[2]	0.9580[3]	3942[3]	3831[3]	3886[3]
	0.9227[2]	4097[2]						
	0.9167[4]	4191[4]						
Experiment	0.9168[6]	4138[6]	1.2746[6]	2991[6]	0.9572[7]	3943[7]	3832[7]	3888[7]

[1] CCSDT/aug-cc-pV5Z, from Ref. [91].
[2] MP4/aug-cc-pVTZ, from Ref. [92].
[3] Diffusion Quantum Monte Carlo, from Ref. [93].
[4] CCSD(T)/aug-cc-pVTZ, from Ref. [77].
[6] From Ref. [94].
[7] From Ref. [95].

acteristic of GGA functionals in general (mean deviations from experiment around 0.01 to 0.02 Å for first- and second-row species [96]) and hence a limitation of the method and not the basis set. The BLYP functional performs worst in this regard. Altmann and co-workers found that BLYP overestimated bond lengths (compared to experiment) for most of the species they studied. [82] Indeed some authors have commented that in some cases, LDA and Hartree-Fock bonds agree more closely with experiment than those from computations using BLYP. [82, 96] In contrast, vibrational frequencies are underpredicted for each of the three species. The BLYP functional performs worst again, especially for HF, for which it is known to be problematic (deviation from experimentally observed fundamental frequency larger than 100 cm^{-1}).

3.4.1 The HF, HCl and H_2O Dimers

Tables 3.2, 3.3 and 3.4 show dissociation energies, geometries and bond length changes for HF, HCl and H_2O dimers.

Geometric predictions for the dimeric species are no better than for the isolated monomers; each functional exaggerates bond lengths. In the case of the numerical basis, bonds are shortened as the size of the basis set is increased. The extent of contraction is significant - in the second decimal place in most cases, to give a qualitative picture. No such effect is observed for the Gaussian sets, with bond lengths staying fairly constant with varying basis set size. Having noted the above, the effect is least pronounced for the PBE functional (the only non-semi-empirical functional included in this study) and bond length changes are most accurately predicted by HCTH.

The Gaussian basis sets underestimate the binding energy of HF dimer (Table 3.2), with or without CP corrections, except in combination with the PBE functional, for which the binding is artificially strong. The BLYP/DNP combination predicts the binding energy to within 2% (comparing to best *ab initio* uncorrected) while HCTH/aug-cc-pVTZ underestimates the energy, even without CP corrections. Conversely, PBE performs very well for the geometric parameter, with both HCTH and BLYP overestimating it, regardless of the basis set.

The DNP basis performs best for energetic predictions of H_2O dimer (Table 3.3) compared to the Gaussian basis sets, with or without CP corrections, when combined with BLYP and HCTH. However, the PBE functional in combination with all basis sets overestimates the binding. The numerical basis set performs very well for the geometric parameter with all functionals. Within the functionals, both BLYP and HCTH slightly overestimate the geometric parameter (HCTH to a greater extent) while PBE performs best again in this respect. Having said this, bond length *changes* are most accurately predicted by HCTH, as stated above for the monomeric species.

The results for HCl dimer are presented in Table 3.4. The reader will notice that the best *ab initio* binding energies differ somewhat from the experimental binding energy. This is because the dynamics of the HCl dimer interaction are very complex and experimentalists [97] acknowledge that the system deviates significantly from equilibrium geometry. Hence, we include the experimental value for completeness only and restrict our comparisons to the best *ab initio* results.

Both types of basis set show similar behaviour (for both energies and geometries) within a given functional, indicating that DNP performs at least as well as the other basis sets for this system. BLYP performs quite poorly for HCl, underestimating the binding energy in combination with both types of basis set, regardless of whether CP corrections are taken into account. PBE tends to overestimate the binding energy, less so when CP corrections are applied. Interestingly, while HCTH slightly underestimates the binding energy, addition of CP corrections only leads to further underestimation. PBE performs best again for the geometry, followed by BLYP.

Table 3.2: Selected geometric and energetic parameters for HF dimer in C_s symmetry. All energies (including BSSE) in kJ mol^{-1}, all bond lengths and bond length differences in Å. The quantity, $\Delta r_{H-F}(d)$, is the difference in the donor bond length when measured in isolation (relaxed) and in the dimer (relaxed).

Functional/Basis	$\mathbf{D}_e$	**BSSE**	$\mathbf{D}_e$**[1]**	$\mathbf{r}_{F\cdots F}$	$\Delta\mathbf{r}_{H-F}$**(d)**
BLYP					
DNP	-19.25		-19.25	2.760	0.0080
6-311+G(3df,2pd)	-17.27	1.630	-18.90	2.763	0.0002
aug-cc-pVTZ	-17.37	0.290	-17.67	2.760	0.0090
HCTH					
DNP	-18.27		-18.27	2.872	0.0050
6-311+G(3df,2pd)	-16.62	1.559	-18.18	2.870	0.0057
aug-cc-pVTZ	-16.57	0.287	-16.86	2.884	0.0053
PBE					
DNP	-21.90		-21.90	2.719	0.0090
6-311+G(3df,2pd)	-20.04	1.794	-21.83	2.725	0.0100
aug-cc-pVTZ	-20.53	0.336	-20.57	2.723	0.0099
Best *ab initio*	-19.1[2]		-19.0[2]	2.735[3]	
	-18.77[4]		-19.77[4]	2.732[4]	0.0056[4]
Experiment	-19.37[5]			2.720[6]	

[1] No counterpoise correction applied.
[2] CCSD(T)/aug-cc-pVQZ, from Ref. [98].
[3] Derived from spectroscopic data, from Ref. [99].
[4] CCSD(T)/aug-cc-pVQZ, from Ref. [91].
[5] Experimental binding energy derived from Ref. [100] with zero point corrections from Ref. [101] and Ref. [102].
[6] Experimental equilibrium geometries from Ref. [103] and Ref. [104].

Table 3.3: Selected geometric and energetic parameters for H_2O dimer in C_s symmetry.

Functional/Basis	D_e	BSSE	D_e[1]	$r_{O\cdots O}$	Δr_{O-H}(d)
BLYP					
DNP	-20.54		-20.54	2.952	0.0080
6-311+G(3df,2pd)	-17.46	1.285	-18.75	2.949	0.0090
aug-cc-pVTZ	-17.46	0.190	-17.65	2.946	0.0080
HCTH					
DNP	-20.27		-20.27	3.033	0.0050
6-311+G(3df,2pd)	-17.38	1.335	-18.72	3.052	0.0063
aug-cc-pVTZ	-17.34	0.174	-17.51	3.062	0.0050
PBE					
DNP	-24.03		-24.03	2.922	0.0080
6-311+G(3df,2pd)	-21.34	1.508	-22.85	2.896	0.0111
aug-cc-pVTZ	-21.43	0.238	-21.67	2.896	0.0100
Best *ab initio*			-21.00[2]	2.912[2]	0.0063[2]
Experiment	-22.59[3]			2.98[3]	

[1] No counterpoise correction applied.
[2] All-electron Coupled Cluster combined with an extrapolation procedure, from Ref. [105].
[3] From Ref. [106] and Ref. [107].

3.4.2 The CO + HF Complexes

The situation is less clear for the CO + HF complex, the results for which are shown in Tables 3.5 and 3.6. Two different bonding modes exist for these molecules: the HF molecule can bond to either the carbon atom or the oxygen atom of CO with the OC(HF) binding mode being the more energetically stable.

Binding energies for the CO(HF) species are less accurately predicted by the DNP set than either of the two Gaussian sets. Within each functional, DNP produces energies which are higher (stronger binding) than those produced by the Gaussian sets. For the BLYP functional, this results in a value which is closer to the best *ab initio* result but which is still far from accurate. The O$\cdots$F bond length is overestimated by all basis sets, most significantly within the HCTH functional. Within the other two functionals, the DNP result is only slightly overestimated compared to the Gaussian sets. The PBE functional reproduces bond length changes fairly accurately and also gives the best estimate for the bond length itself.

Table 3.4: Selected geometric and energetic parameters for HCl dimer in C_s symmetry.

Functional/Basis	$\mathbf{D}_e$	**BSSE**	$\mathbf{D}_e$[1]	$\mathbf{r}_{Cl\cdots Cl}$	$\Delta\mathbf{r}_{H-Cl}$**(d)**
BLYP					
DNP	-5.33		-5.33	3.891	0.0060
6-311+G(3df,2pd)	-4.10	1.662	-5.76	3.858	0.0068
aug-cc-pVTZ	-4.61	0.253	-4.863	3.88	0.0066
HCTH					
DNP	-8.05		-8.05	4.014	0.0040
6-311+G(3df,2pd)	-6.87	1.295	-8.17	4.026	0.0024
aug-cc-pVTZ	-7.28	0.115	-7.39	4.055	0.0031
PBE					
DNP	-9.32		-9.32	3.675	0.0090
6-311+G(3df,2pd)	-8.13	2.019	-10.15	3.708	0.0092
aug-cc-pVTZ	-8.71	0.257	-8.97	3.737	0.0099
Best *ab initio*	-7.75[2]	0.75[2]	-8.5[2]	3.746[2]	
	-8.27[3]			3.746[3]	
Experiment	-5.24[4]				

[1] No counterpoise correction applied.
[2] CCSD(T)/aug-cc-pVQZ, from Ref. [28].
[3] Determined from a PES constructed from far-infrared and near-infrared spectroscopic measurements, from Ref. [108].
[4] From Ref. [97].

Table 3.5: Selected geometric and energetic parameters for CO(HF) complex in C_{2v} symmetry.

Functional/Basis	D_e	BSSE	D_e[1]	$r_{O\cdots F}$	Δr_{H-F}(d)
BLYP					
DNP	-5.86		-5.86	3.050	0.0010
6-311+G(3df,2pd)	-4.70	1.288	-5.99	3.016	0.0009
aug-cc-pVTZ	-4.63	0.452	-5.08	3.034	0.0028
HCTH					
DNP	-7.72		-7.72	3.198	0.0010
6-311+G(3df,2pd)	-6.45	1.115	-7.57	3.122	0.0017
aug-cc-pVTZ	-6.52	0.509	-7.03	3.206	0.003
PBE					
DNP	-7.57		-7.57	3.009	0.0020
6-311+G(3df,2pd)	-6.34	1.458	-7.79	2.980	0.0028
aug-cc-pVTZ	-6.39	0.492	-6.88	3.001	0.0027
Best *ab initio*[2]	-6.99	1.92	-8.91	2.9517	0.0020

[1] No counterpoise correction applied.
[2] CCSD(T)/aug-cc-pVQZ, from Ref. [28].

For the OC(HF) complex, the DNP basis produces the most notably different energy results in comparison to the Gaussian basis sets (in combination with each of the three functionals) while the geometry is surprisingly well predicted. We note that Tuma and co-workers obtained binding energies for this system which spanned a range of 10 kJ for the functionals being tested. Although our own binding energies span a range of 12 kJ (uncorrected), most of the range can be attributed to the DNP results. It should also be noted that within the PBE functional, the Gaussian-type basis sets also produce unusally strong binding energies. The cause of this spurious overbinding is unknown, though it does seem to suggest that it is not the weakness of the interaction (remembering that CO(HF) is more weakly bound yet better behaved) but the different nature of the hydrogen bond which may be important. We also point out that the best *ab initio* value quoted both in this study and the work of Tuma [28] is the only high-level value available in the literature for comparison. Further investigation with a different high-level method (such as QMC) may shed light on the unusual behaviour of this system.

Table 3.6: Selected geometric and energetic parameters for OC(HF) complex in C_{2v} symmetry.

Functional/Basis	D_e	BSSE	D_e [1]	$r_{C\cdots F}$	Δr_{H-F}
BLYP					
DNP	-24.35		-24.35	2.977	0.0110
6-311+G(3df,2pd)	-14.72	1.489	-16.21	2.963	0.0102
aug-cc-pVTZ	-14.94	0.611	-15.55	2.971	0.0103
HCTH					
DNP	-22.43		-22.43	3.084	0.0070
6-311+G(3df,2pd)	-14.35	1.429	-15.78	3.046	0.0073
aug-cc-pVTZ	-14.66	0.538	-15.19	3.051	0.0074
PBE					
DNP	-28.27		-28.27	2.896	0.0130
6-311+G(3df,2pd)	-19.04	1.727	-20.77	2.899	0.0128
aug-cc-pVTZ	-19.42	0.631	-20.05	2.905	0.0127
Best *ab initio* [2]	-14.62	2.02	-16.64	2.987	0.0066

[1] No counterpoise correction applied.
[2] CCSD(T)/aug-cc-pVQZ, from Ref. [28].

3.4.3 Phosphinic Acid Dimer

Phosphinic acid (PA) dimer, $(H_2PO(OH))_2$, is the largest complex considered in this study and also most closely resembles our system of interest, phenylphosphonic acid dimer; the structure of the dimer is shown in Figure 3.1. A previous study by Gonzalez and co-workers [109] involved the calculation of binding energies, geometries and harmonic vibrational frequencies for both phosphinic acid dimer and its dimethyl derivative. They found that, quite surprisingly, the binding energy calculated for PA dimer is equal to that of dimethyl phosphinic acid (DMPA) and is in good agreement with the only experimental value for DMPA (5.7 ± 1 kcal mol^{-1}). Due to the large size of the system, Coupled Cluster calculations of any sort cannot be performed and so we are forced to be content with a relative comparison only.

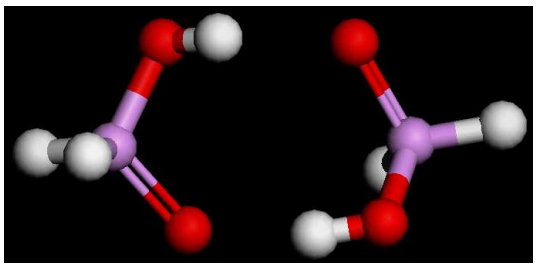

Figure 3.1: Structure of the phosphinic acid dimer. Different atoms are distinguished by the following colours: purple–phosphorus, red–oxygen, white–hydrogen.

The binding energy obtained with the DNP basis is closest to the best available literature value and performs best in combination with HCTH. Within the other two functionals, DNP produces results which are very close to the Gaussian basis sets, albeit the energies are consistently overestimated. With respect to the geometric parameter, DNP produces results consistent with the Gaussian basis sets within each functional, with PBE producing the most accurate result. Based on these findings, it would be reasonable to expect the numerical basis sets to give satisfactory results when applied to phenylphosphonic acid dimer.

Table 3.7: Selected geometric and energetic parameters for phosphinic acid dimer.

Functional/Basis	$\mathbf{D}_e$	**BSSE**	$\mathbf{D}_e$[1]	$\mathbf{r}_{O\cdots O}$	$\Delta\mathbf{r}_{O-H}$
BLYP					
DNP	-6.31		-6.31	2.641	0.045
6-311+G(3df,2pd)	-6.35	0.19	-6.54	2.624	0.051
aug-cc-pVTZ	-6.28	0.07	-6.35	2.617	0.050
HCTH					
DNP	-5.55		-5.55	2.660	0.037
6-311+G(3df,2pd)	-5.40	0.23	-5.63	2.630	0.042
aug-cc-pVTZ	-5.47	0.08	-5.55	2.633	0.041
PBE					
DNP	-7.54		-7.54	2.577	0.059
6-311+G(3df,2pd)	-7.65	0.23	-7.88	2.552	0.065
aug-cc-pVTZ	-7.65	0.08	-7.72	2.558	0.0648
Best available[2]	-5.7			2.407	

[1] No counterpoise correction applied.
[2] B3LYP/6-311+G(3df,2pd), from Ref. [109].

3.5 Error Analysis

Tables 3.8 and 3.9 show the mean difference between a particular set of values (either geometries or energies) and the best *ab initio* value and also the mean deviation, where the two quantities have been calculated thus:

$$\overline{x} = \frac{\sum_{i=1}^{N} |x_i - y|}{N} \tag{3.2}$$

where $\overline{x}$ denotes the mean, N is the number of observations, x_i is the value of the ith observation and y is the best *ab initio* result. We have adopted a conservative approach and have assumed that the errors are additive (hence, we have calculated the absolute difference), or rather, we have not assumed that the errors cancel in any way (though of course, they may). Hence, all values quoted reflect the 'worst-case scenario'.
The mean deviation, *MD*, is formally defined [110] as,

$$MD = \frac{1}{N}\sum_{i=1}^{N}|x_i - \overline{x}| \tag{3.3}$$

The mean deviation then provides a measure of how uniformly these 'errors' occur, *i.e.* the closer x_i is to $\overline{x}$, the closer the expression $|x_i - \overline{x}|$ will be to zero and consequently, the lower the value of *MD*. Low values of *MD* indicate that errors deviate only slightly from the mean and as such *MD* provides a measure of stability of each method.

From Tables 3.8 and 3.9 it can be seen that in terms of energies, DNP shows the least consistency within each functional compared to the Gaussian basis sets. However, with respect to geometries, it is at least as consistent as the Gaussian basis sets. The HCTH/6-311+G(3df,2pd) combination seems to produce the lowest mean and mean deviations with HCTH performing very well with both the DNP and aug-cc-pVTZ bases. However, HCTH has consistently overestimated the geometric parameter for the systems considered in this study, and this is reflected in an increased mean error (Table 3.9), though geometries have been *consistently* overestimated, as indicated by a reasonably low mean deviation. Conversely, the PBE functional produces the highest mean and mean deviations for energies but seems to be the most reliable for geometries, producing the lowest mean/mean deviation.

Table 3.8: Deviations from best *ab initio* (uncorrected) binding energies. All entries in kJ mol^{-1}.

Functional/Basis	Mean	Mean Deviation
BLYP		
DNP	2.93	2.06
6-311+G(3df,2pd)	1.69	1.14
aug-cc-pVTZ	2.65	1.15
HCTH		
DNP	1.78	0.80
6-311+G(3df,2pd)	0.73	0.31
aug-cc-pVTZ	2.01	0.53
PBE		
DNP	3.94	3.07
6-311+G(3df,2pd)	2.31	0.93
aug-cc-pVTZ	1.63	0.87
OVERALL		
DNP	2.88	2.28
6-311+G(3df,2pd)	1.58	0.98
aug-cc-pVTZ	2.10	0.97

We conclude this study with a brief analysis of (relative) CPU timings for geometry optimisation of phosphinic acid dimer. Table 3.10 shows the relative time taken for optimisation for each functional and basis set combination. All times have been normalised as a percentage to the slowest combination of functional/basis set, BLYP/aug-cc-pVTZ, which has been set to 100%. Whilst this analysis is machine-dependent, we believe the trends are strong enough that they could be qualitatively reproduced on a different machine.

The DNP basis is some 10 times faster at worst (comparing slowest numerical to fastest GTO) and 116 times faster at best (comparing fastest numerical to slowest GTO). Within the Gaussian basis sets, the BLYP/aug-cc-pVTZ combination is slowest (as stated above) and the HCTH/6-311+G(3df,2pd) is fastest. In general, the aug-cc-pVTZ set is slower than the 6-311+G(3df,2pd) set. Considering that both types of basis set give similar geometries, it should be possible to perform reasonably accurate geometry optimisations on large systems using numerical basis sets.

Table 3.9: Deviations from best *ab initio* geometries. All entries in Å.

Functional/Basis	Mean	Mean Deviation
BLYP		
DNP	0.064	0.046
6-311+G(3df,2pd)	0.053	0.028
aug-cc-pVTZ	0.060	0.040
HCTH		
DNP	0.174	0.066
6-311+G(3df,2pd)	0.157	0.054
aug-cc-pVTZ	0.185	0.077
PBE		
DNP	0.048	0.030
6-311+G(3df,2pd)	0.034	0.023
aug-cc-pVTZ	0.034	0.026
OVERALL		
DNP	0.096	0.059
6-311+G(3df,2pd)	0.082	0.058
aug-cc-pVTZ	0.093	0.071

Table 3.10: Comparison of relative times taken to optimise the geometry of phosphinic acid dimer.

Functional	Basis	Relative Time
BLYP	DNP	0.93
	6-311+G(3df,2pd)	22.00
	aug-cc-pVTZ	100.0
HCTH	DNP	0.85
	6-311+G(3df,2pd)	14.81
	aug-cc-pVTZ	62.39
PBE	DNP	1.23
	6-311+G(3df,2pd)	13.31
	aug-cc-pVTZ	68.87

3.6 Conclusions

In general, the numerical basis sets produce stronger binding energies than the Gaussian bases (a conclusion also reached by Altmann and co-workers [82] in their study on sulfur-containing molecules) while producing slightly higher mean deviations for energies.

The numerical basis sets are comparable (sometimes better) to the Gaussian bases in terms of geometries. We have also compared the performance of the GGA functionals employed in this study and found that for the systems considered here the BLYP functional performs the worst, both in terms of energies and geometries. HCTH tends to consistently overestimate geometries, but produces reasonably accurate energies. PBE performs very well for geometries but consistently overestimates binding energies, in agreement with Tuma's study. [28]

Overall, numerical basis sets offer a definitive advantage over Gaussian bases where geometry optimisation of large systems is concerned. Given that our systems of interest are quite large (tens of atoms, including heavy elements), the amount of computational effort that would need to be expended if one were using a Gaussian basis set would not be commensurate with a significant gain in accuracy, especially when geometry is the primary property of interest. However, as with any technique, caution should be exercised when the systems of interest are weakly or unusually bound.

Chapter 4

Hydrogen Bonding Interactions in the Phenylphosphonic Acid Dimers

4.1 Introduction

In the previous chapter we tested the performance of numerical basis sets in hydrogen bonding applications and compared the results with those from calculations using Gaussian basis sets. The numerical basis sets gave accurate geometries at much lower computational cost but had a slight tendency to overbind. In this chapter, we use numerical basis sets to study the hydrogen bonding interaction between phenylphosphonic acid dimers in the organophosphonates.

The phenylphosphonic acid dimers are actually anions formed from the interaction between phenylphosphonic acid and a phenylphosphonate anion. The resulting hydrogen bond is thus ionic. Like the conventional hydrogen bond, ionic hydrogen bonds are involved in a broad range of phenomena: ionic clusters and nucleation, in the structures of ionic crystals, surfaces, silicates and clays and in self-assembly in supramolecular chemistry and molecular crystals. They are also important in biological processes, such

as protein folding, formation of membranes and biomolecular recognition. A fundamental understanding of the properties of ionic hydrogen bonds is thus vital to the understanding of many different phenomena.

The primary interaction in the phenylphosphonic acid dimers is an O-H-O^- bond and it is therefore an anionic hydrogen bond. Ionic hydrogen bonds are typically characterised by short bond lengths and strong binding energies in the range -(5-35) kcal/mol. [111] Anionic O-H-O^- interactions are no exception with bond lengths in the range 2.4-2.5 Å and binding energies of about -28 kcal/mol. [112, 113, 114] As a point of comparison, the O-O distance in water dimer, classified as a neutral moderate hydrogen bond, is 2.976 Å with a binding energy of about -5 kcal/mol. Short hydrogen bonds, such as O-H-O^- bonds, are usually centrosymmetric, which means that the proton is drawn to the midpoint of the O-O axis and therefore can no longer be associated with either a donor or acceptor. This has implications for the theoretical treatment of such bonds, as we will discuss below.

Phenylphosphonic acid is one of the most common phosphonic acids used to prepare metal phosphonates comprising layered structures. Phosphonic acids are commonly observed to form short, linear hydrogen bonds as the basis of their solid state structures. However, quantitative data (both theoretical and experimental) as to the strength of these bonds is lacking. In this chapter, we use theoretical methods to study both the geometry and energetics of the hydrogen bonding interaction in the organophosphonates. The effect of substituents on the strength and geometry of the hydrogen bonding interaction is also considered. Further, we aim to understand, and possibly quantify, the effect of the crystal field on the structure of the phenylphosphonic acid molecules in the organophosphonates.

4.2 Previous Studies

One of the most extensively studied anionic hydrogen-bonded systems has been the hydroxide hydrate anion, $(HOH\cdots OH)^-$. An initial discrepancy in the experimentally obtained enthalpy of formation sparked several theoretical studies aimed at determining the structure and binding energy. Early studies by Newton and Ehrenson [115]

at the HF/4-31G level of theory suggested a binding energy of -40.7 kcal/mol. The minimum energy structure was found to have an O$\cdots$O distance of 2.597 Å with a flat potential energy surface for transfer of the proton from donor to acceptor, a feature characteristic of short, particularly centrosymmetric, hydrogen-bonded systems. Energy calculations performed by Del Bene [116] at the MP4SDQ/6-31+G(2d,2p) level of theory on Hartree-Fock geometries yielded a much lower binding energy of -26.6 kcal/mol. This study also emphasised the importance of including electron correlation and adding diffuse functions to the basis on non-hydrogen atoms. Gronert [112] calculated the potential energy surface for proton transfer using MP2/6-31+G(d,p) geometries and located a transition state with a binding energy of -27.8 kcal/mol. The transition state disappeared with the inclusion of zero-point energy corrections. DFT calculations [113] predict an O$\cdots$O distance of 2.463 Å, a binding energy of -28.62 kcal/mol and a 0.1 kcal/mol barrier to proton transfer. [114] These findings suggest that the amplitude for proton motion along the O$\cdots$O axis is large.

There have also been studies concerning the effect of substituents on the binding energy and shape of the proton transfer potential. Wolfe, *et al* [117] calculated the proton transfer potential for the $(CH_3OH\cdots OCH_3)^-$ system at both the HF and MP levels of theory. The binding energy, calculated at the MP2/6-31+G(d,p) level of theory on HF geometries, was determined to be -27.4 kcal/mol. This was consistent with experimental measurements, which indicated that methyl substitution increased the binding by 1-2 kcal/mol. The HF calculations also predicted a small barrier to proton transfer, which subsequently disappeared with the inclusion of electron correlation. Still larger substituents appear to have little effect on the binding energy; the binding energy of acetate dimer, $(CH_3COOH\cdots OOCH_3)$, was determined experimentally [118] to be -29.3 kcal/mol, less than 1 kcal/mol lower than the lowest energy calculated for hydroxide hydrate anion.

The dimer of phosphinic acid is currently the strongest known neutral dimer and also closely resembles our system of interest, phenylphosphonic acid dimer. DFT calculations [109] have predicted a binding energy of -23.2 kcal/mol, which is the same as that calculated for the dimethyl derivative of phosphinic acid and in good agreement with

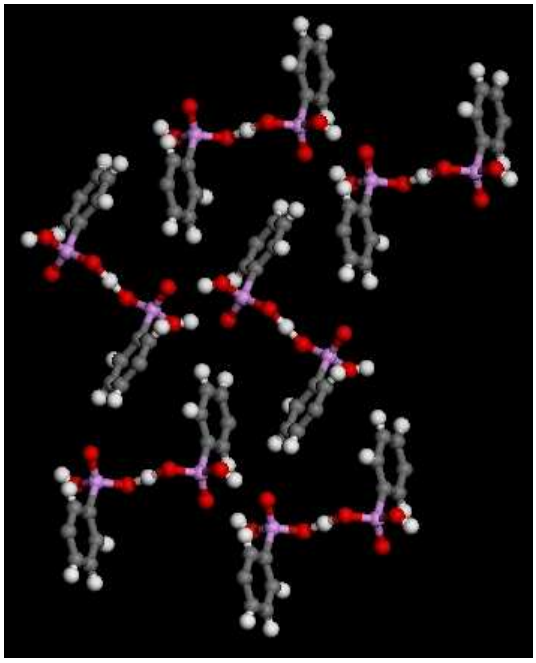

Figure 4.1: Hydrogen bonding interaction between phenylphosphonic acid dimers.

the only experimental measurement available (-23.9 ± 6 kcal/mol). These findings are consistent with those for hydroxy hydrate anion, namely that substituents appear to have a minor effect on the strength of the hydrogen bonding interaction.

4.3 Computational Details

4.3.1 System, Functional and Basis Set Selection

In Chapter 1 we described the structure and composition of the organophosphonates under study in this work. We noted that each member of the family contains chains of phenylphosphonic acid molecules linked by single, linear hydrogen bonds to form dimers. These negatively charged dimers form two hydrogen bonds with each other to form chains along the b axis of the crystal. There are also dispersion interactions between the aromatic rings of dimers in adjacent chains to form sheets. A fragment of one of these sheets is shown in Figure 4.1. We have thus simulated phenylphosphonic acid dimer in two different hydrogen bonding scenarios to represent the two arrangements present in the molecular crystal: a negatively charged dimer linked by a single

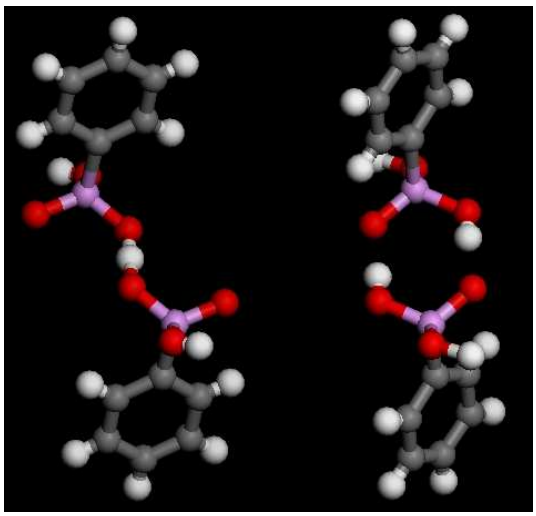

Figure 4.2: Phenylphosphonic acid dimer model systems. Left: single, linear H-bonded dimer. Right: two H-bonds dimer.

hydrogen bond and a neutral dimer of acid molecules linked by two hydrogen bonds, as shown in Figure 4.2. The initial coordinates for each structure were obtained from the bulk crystal, the structure of which was determined by single-crystal x-ray diffraction. [3] The geometries of each model system were optimised, without constraints and the binding energy was calculated as follows:

$$\Delta E = E_{AB} - (E_A + E_B) \tag{4.1}$$

where E_{AB} is the total energy of the relaxed dimer and E_A and E_B are the total energies of the relaxed constituent fragments. Since binding energies of stable complexes are always negative, a higher binding energy denotes a less negative energy and hence less strong binding. Counterpoise corrections were not performed in this work. As stated in the previous chapter, this procedure is not implemented in DMol^3.

One conclusion from the numerical basis set study of the previous chapter was that the PBE functional performed very well for geometries but had a tendency to overbind. In contrast, the HCTH and BLYP functionals performed less well for geometries but produced more accurate energies. These trends were reflected in the results for phosphinic acid dimer, the test system which most closely resembles phenylphosphonic acid dimer. Since the only experimental information available for phenylphosphonic acid

dimer is the geometric parameters, it would make sense to choose a functional on the basis of its ability to accurately predict geometries. Hence, the geometry of each model system has been optimised and the binding energy calculated with the PBE functional in combination with the DNP basis. We have assumed, based on the data for phosphinic acid dimer, that PBE will produce a lower binding energy (too strong binding) for phenylphosphonic acid dimer than the HCTH or BLYP functionals. To test this assumption, we have also performed single-point energy calculations using the HCTH functional on the PBE-optimised geometry for the single H-bond dimer only.
No ionic hydrogen-bonded dimers were included in the numerical basis set study. To confirm the reliability of the numerical basis sets for these systems, we have also calculated the binding energy of the single H-bond dimer using the PBE functional in combination with the Gaussian-type 6-31+G(d,p) basis set.

4.4 Results and Discussion

4.4.1 Relaxations and Binding Energy – Unsubstituted Dimers

The energy of hydrogen bonds in the solid state cannot be directly measured. However, computer simulations allow us to isolate these interactions and study them quantitatively in the gas phase. Geometry and continuum effects tend to decrease the strengths of ionic hydrogen bonds in solids so gas phase binding energies provide an upper limit to the bond strength. [111]

The binding energy for the optimised single H-bond arrangement was calculated to be -27.1 kcal/mol with the PBE functional and DNP basis and -26.4 kcal/mol with the Gaussian-type basis, thus confirming the suitability of the numerical basis sets to describe ionic hydrogen-bonded systems. An additional calculation on the PBE-optimised geometry with the HCTH functional and DNP basis yielded a binding energy of -23.6 kcal/mol, which suggests that the energies obtained with the PBE functional are slightly over-estimated. All energies obtained are consistent with those calculated for phosphinic acid dimer and for the other $(HOH\cdots OH)^-$ systems, described earlier.

The geometry of the PBE-optimised structure is in excellent agreement with the ex-

Table 4.1: Comparison of main geometric parameters between experimental and optimised (PBE/DNP) structures for single H-bond dimer.

	Experiment	**Optimised**
d(O $\cdots$ O), Å	2.426	2.414
d(O $\cdots$ H), Å	1.213	1.207
O-H-O,^{o}	180	180
C-P-O-H,^{o}	82.250	62.663

perimental one, as shown in Table 4.1. The biggest discrepancy concerns the C-P-O-H torsion angle, which is larger in the experimental structure than the optimised one and is probably due to compression effects in the crystal. Crystal field effects will be discussed in more detail in Section 4.4.3.

Calculations were also performed on the unrelaxed structure, *i.e.* the dimer was held fixed at the experimental geometry. The binding energy was computed to be -54.7 kcal/mol with the DNP basis and -54.2 kcal/mol with the Gaussian-type basis. The difference in binding between the relaxed and experimental structures may be understood in terms of the O-O distance. The O-O distance in the relaxed dimer is shorter than in the experimental one, hence we might expect the optimised structure to be more strongly bound. However, this shortening pushes the two negative oxygen atoms closer together, thereby increasing repulsion and slightly de-stabilising the dimer. A higher binding energy (less negative energy, therefore less strong binding) is the final result.

We have also performed calculations on the uncharged double H-bond dimer. The main geometric parameters for both the optimised and experimental structure are compared in Table 4.2. The binding energy of the relaxed structure is calculated to be -22.8 kcal/mol, slightly weaker than that obtained for the single H-bond dimer, as expected because the double H-bond dimer is not charged. The binding energy of the dimer at the experimental geometry was computed to be -15.4 kcal/mol. Table 4.2 reveals that the O $\cdots$ H distance in the optimised structure is shorter than in the experimental one, though the O $\cdots$ O is slightly longer. It is reasonable to infer that this shortening results in an increased binding energy for the optimised structure. This is in contrast to the single H-bond dimer where shortening of the O $\cdots$ O distance in the optimised

structure lead to increased repulsion between the oxygens and hence a higher binding energy than the experimental structure.

Table 4.2: Comparison of main geometric parameters between experimental and optimised (PBE/DNP) structure for double H-bond dimer.

	Experiment	**Optimised**
d(O···O), Å	2.517	2.572
d(O···H), Å	1.724	1.546
O-H-O,°	175.193	175.129
C-P-O-H,°	102.888	95.991

Insight into the nature of the hydrogen bonding interaction between phenylphophonic acid molecules may be obtained by examining the electron density map of the single H-bond dimer, Figure 4.3. Positive regions (red) correspond to electron deficient areas while negative regions (blue) correspond to electron rich areas. We note that although electron density maps are quantitatively calculated, their real value lies in their use as a tool to qualitatively compare the electron distributions in chemically different structures.

Figure 4.3 shows that much of the electron density is concentrated in and around the hydrogen bond which links the phenylphosphonic acid and anion molecules. This area contains a number of oxygen atoms, which pull electron density out of the aromatic rings thereby leaving them electron deficient. Addition of a second dimer to form a chain fragment does little to alter this behaviour, as shown in Figure 4.4. The hydrogen bonds linking the phenylphosphonic acid molecules form an electron-rich 'spine', which runs along the b axis of the crystal.

4.4.2 Relaxations and Binding Energy – Substituted Dimers

A second series of mixed-ligand copper organophosphonate materials was recently synthesised. [119] These new compounds are similar to the materials described in Chapter 1, but phenylphosphonic acid was replaced by a substituted phenylphosphonic acid derivative, *i.e.* the aromatic ring bears either a nitro (NO_2) group or methyl (CH_3) group in place of one hydrogen, as shown in Figure 4.5 for the nitro substituted dimer. The molecular unit of the new structures is analogous to that of the parent compounds – a five-coordinate $[Cu(phen)_2X]^+$ cation with two acid molecules in some arrangement functioning as a counter-ion. In this section, we will consider the effects of the ni-

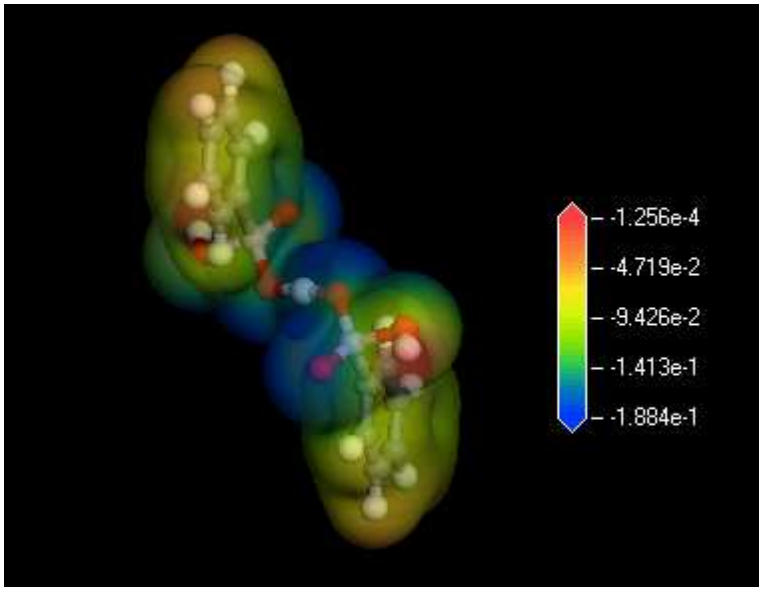

Figure 4.3: Electrostatic potential mapped onto the electron density calculated at the PBE/DNP level of theory for the single H-bond dimer.

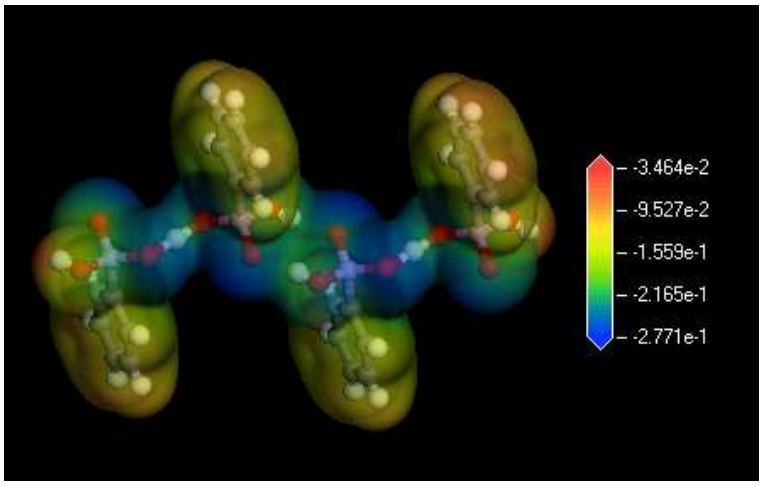

Figure 4.4: Electrostatic potential mapped onto the electron density calculated at the PBE/DNP level of theory for two single H-bond dimers.

Table 4.3: Comparison of main geometric parameters (experimental) between nitro-substituted and unsubstituted phenylphosphonic acid dimers.

	Nitro-substituted	**Unsubtituted**
d(O···O), Å	2.433	2.426
d(O···H), Å	1.498	1.213
d(O···H), Å	0.943	1.213
O-H-O,°	170.748	180.000
C-P-O-H,°	74.921	82.250
C-P-O-H,°	79.243	82.250

tro and methyl substituents on the geometry and energetics of the hydrogen bonding interaction.

The molecular unit of the nitro-substituted organophosphonate contains an acid and anion molecule linked by a single H-bond. The main geometric parameters of the nitro substituted dimer are compared with those of the unsubstituted parent in Table 4.3. An immediate observation is that the nitro-substituted dimer is not symmetric; the two torsion angles are inequivalent as are the O···H bond lengths. It follows from the latter observation that the hydrogen bond in the nitro-substituted dimer is not centrosymmetric, *i.e.* the hydrogen atom can be unambiguously associated with the proton-donating (acid) fragment.

The geometry of the nitro-substituted dimer was optimised in the same manner as that for the unsubstituted dimers, with the PBE functional and DNP basis set. Table 4.4 shows that the optimised structure is more symmetric than the experimental one. The O···O distance is shorter in the optimised structure and the hydrogen atom has been drawn away from the donating acid towards the midpoint of the hydrogen bond, which

Table 4.4: Comparison of main geometric parameters between experimental and optimised (PBE/DNP) structure for nitro-substituted phenylphosphonic acid dimer.

	Experiment	**Optimised**
d(O···O), Å	2.433	2.422
d(O···H), Å	1.498	1.154
d(O···H), Å	0.943	1.268
O-H-O,°	170.748	178.741
C-P-O-H,°	74.921	60.355
C-P-O-H,°	79.243	58.265

is considerably more linear (178.741°) than in the experimental structure (170.748°). The largest discrepancy between experiment and theory again concerns the C-P-O-H torsion angles. A contraction of the O···O distance and a decrease in C-P-O-H torsion angle was also observed for the unsubstituted single H-bond dimer.

We have also examined the effect of the substituent on the binding energy of the dimer. Nitro groups are known to be strong electron withdrawing agents hence we might expect the hydrogen bond in the nitro-substituted dimer to be weaker than its unsubstituted analogue. Furthermore, the O···O distance in the nitro-substituted dimer is slightly longer than in the unsubstituted parent and this also suggests that the hydrogen bond may be weaker. In fact, the binding energy of the nitro-substituted dimer is calculated to be -31.75 kcal/mol, about 5 kcal/mol *stronger* than the unsubstituted single H-bond dimer.

This initially counter-intuitive result may be readily explained by considering the dissociation constants of the substituted and nitro-substituted dimers. Nagarajan, *et al* determined the dissociation constants of 36 substituted phenylphosphonic acids. [120] They found that nitro-substituted phenylphosphonic acid had a significantly lower dissociation constant (pKa=1.28) than its unsubstituted parent (1.86). The dissociation constant is a measure of how readily an acid dissociates into H^+ and its conjugate base, i.e. how readily it will donate its proton (for a monoprotic acid); lower dissociation constants denote stronger acids. Since the formation of a hydrogen bond weakens the A-H bond of the proton donor, we can argue that acids which readily donate their protons (strong acids) may more readily form hydrogen bonds. Hence, we could argue that nitro-substituted phenylphosphonic acid is a stronger acid and we could therefore expect it to form a stronger hydrogen bond. Another possibility is that the stronger binding is merely a result of geometric differences between the substituted and unsubstituted single H-bond dimers. To test this hypothesis, we performed an additional single-point energy calculation on the relaxed nitro-substituted dimer whereby the nitro groups were replaced with hydrogen atoms. The binding energy of this structure was calculated to be -26.69 kcal/mol, similar to that obtained for the unsubstituted single H-bond dimer. Hence, the stronger binding energy of the nitro-substituted dimer

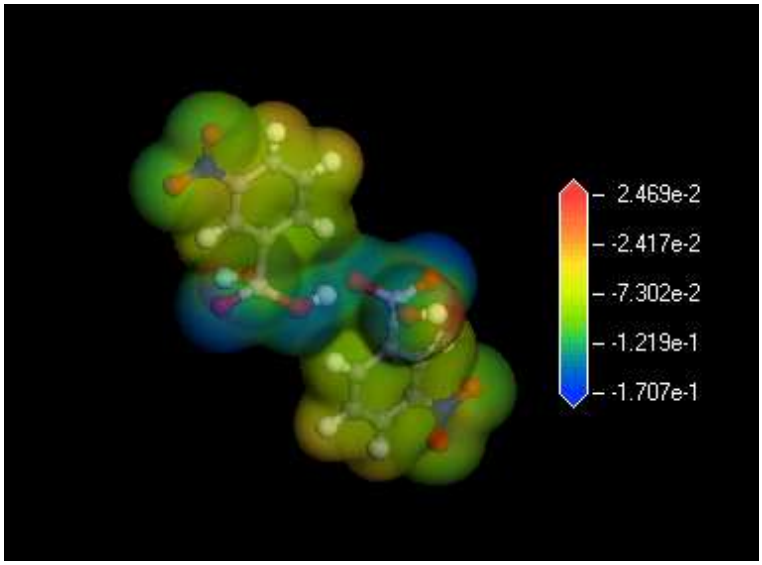

Figure 4.5: Electrostatic potential mapped onto the electron density calculated at the PBE/DNP level of theory for the nitro-substituted single H-bond phenylphosphonic acid dimer.

must have electronic effects (exerted by the nitro groups) at its origin.

Figure 4.5 shows that the electron density distribution of the nitro-substituted dimer is qualitatively similar to that of its unsubstituted parent: much of the electron density is concentrated in and around the hydrogen bond linking the two fragment molecules. The oxygen atoms attached to the nitrogen of the nitro group are also electron-rich.

The molecular unit of the methyl-substituted organophosphonate also contains an acid and anion molecule linked by a single H-bond. The structure of the methyl-substituted single H-bond dimer is almost identical to that of its unsubstituted parent, as Table 4.5 shows. The hydrogen bond linking the two fragments is linear and centrosymmetric. In contrast to both the nitro and unsubstituted dimers, geometry optimisation lengthens, rather than shortens, the O$\cdots$O distance, Table 4.6. Given the similarity between the methyl-substituted and unsubstituted geometries and the fact that methyl groups are only very weakly electron donating, we might expect the binding energy of the methyl-substituted dimer to be very similar to the unsubstituted single H-bond dimer. This is indeed the case; the binding energy of the methyl-substituted dimer is predicted to be -26.01 kcal/mol. Figure 4.6 shows that the electron density distribution of the methyl-substituted dimer is almost identical to that of the unsubstituted dimer and

Table 4.5: Comparison of main geometric parameters (experimental) between methyl-substituted and unsubstituted phenylphosphonic acid dimers.

	Methyl-substituted	**Unsubstituted**
d(O···O), Å	2.410	2.426
d(O···H), Å	1.208	1.213
O-H-O,°	180.000	180.000
C-P-O-H,°	87.978	82.250

Table 4.6: Comparison of main geometric parameters between experimental and optimised (PBE/DNP) structures for methyl-substituted phenylphosphonic acid single H-bond dimer.

	Experiment	**Optimised**
d(O···O), Å	2.410	2.415
d(O···H), Å	1.208	1.208
O-H-O,°	180.000	179.998
C-P-O-H,°	87.978	75.805

very similar to that of the nitro-substituted dimer.

4.4.3 Crystal Field Effects

The crystalline field is the total influence of the crystal surroundings on a target molecule. A large amount of information about molecular structure now comes from the crystal structure. Implicit in these descriptions is the assumption that the crystal environment has not affected bond lengths and therefore the overall structure of the molecule. This hypothesis was originally formulated by the Russian crystallographer Kitaigorodsky [121], who believed that the major effect of the crystalline field on molecular structure was to compress the molecule and promote favourable stacking by distorting torsion angles. In recent studies [122, 123] of crystal structures containing metal complexes, it was found that the crystalline field can indeed significantly affect bond lengths. Let us then examine the hydrogen bonding interaction between the unsubstituted phenylphosphonic acid dimers in light of Kitaigorodsky's hypothesis.

In Sections 4.4.1 and 4.4.2 we saw how geometry optimisation of the single H-bonded dimers tended to decrease C-P-O-H torsion angles and shorten O···O distances. The optimisations were performed in the gas-phase, *i.e.* the dimers were modelled in complete isolation from the crystal environment. Hence, it is pertinent to ask, which

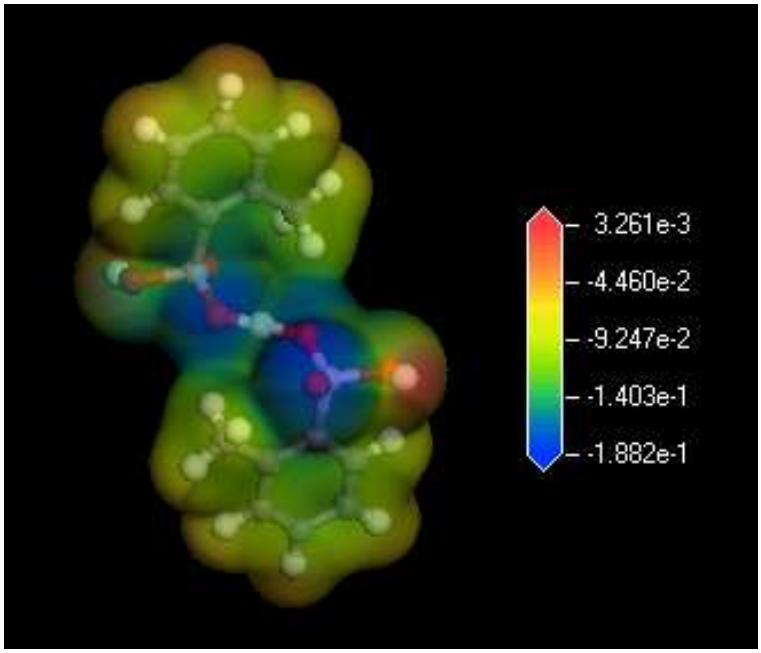

Figure 4.6: Electrostatic potential mapped onto the electron density calculated at the PBE/DNP level of theory for the methyl-substituted single H-bond phenylphosphonic acid dimer.

features of the crystal environment are responsible for the observed experimental geometries of the single H-bonded dimers?

Figure 4.1 shows that in addition to hydrogen bonding interactions between dimers to form chains, there are also dispersion interactions between the aromatic rings of the dimers to form an acid 'sheet'. As a first step, let us assume that the discrepancies between the experimental and optimised structures are caused by a failure to account for these aromatic ring interactions. Interactions between aromatic rings are very weak, hence the rings probably represent more of a steric, rather than chemical constraint. To test our assumption, we have constructed a model of the single H-bond dimer in which the aromatic rings of adjacent dimers have been included, as shown in Figure 4.7. The aromatic rings of the adjacent dimers were constrained in their experimental positions while the single H-bond dimer was allowed to relax, without constraints. In this model, the O$\cdots$O distance still shortens to 2.417 Å but the C-P-O-H angle only decreases to 74.989^{o}, in better agreement with the experimentally observed angle of 82.250^{o}. Hence, steric constraints imposed by the aromatic rings are not alone responsible for the experimentally observed geometry.

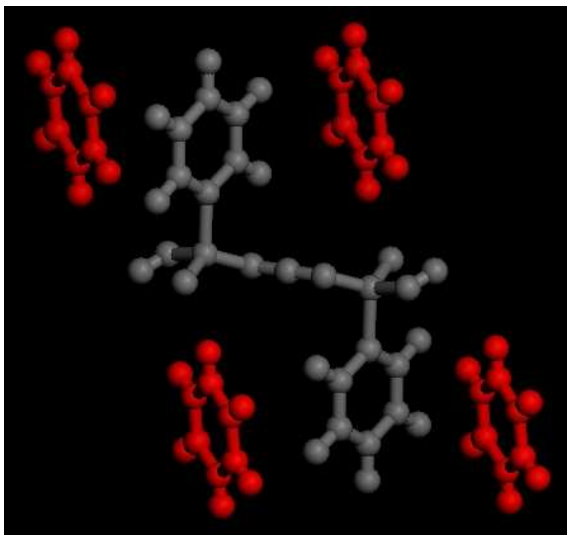

Figure 4.7: Unsubstituted single H-bond phenylphosphonic acid dimer with aromatic rings from adjacent dimer molecules. Red atoms have been constrained in their experimental positions while grey atoms have been allowed to relax, without constraints.

We have made repeated mention of the hydrogen bonding interaction between dimers to form chains. These hydrogen bonds could be considered as chemical rather than steric constraints. It would thus be sensible to construct a model in which these interactions are included. Hence, in addition to the aromatic rings of adjacent dimers, we have added two single H-bond dimers which form hydrogen bonds with our target dimer, as shown in Figure 4.8. Once again, all molecules except for the target dimer were constrained in their experimental positions while the target dimer was allowed to relax, without constraints. In this model, the O $\cdots$ O distance still contracts to 2.418 Å but the C-P-O-H angle *increases* to 109.681°. Steric constraints together with chemical constraints imposed by the hydrogen bonding interaction are still not sufficient to reproduce the experimentally observed geometry. Furthermore, a possible problem with this model is that it carries a large negative charge of -3 (each dimer unit carries a charge of -1). In the crystal, the phenylphosphonic acid dimer anions are charge-balanced by the copper-halogen cations, which carry a charge of +1.

In the final model for this section, we include the charge-balancing cation in addition to the aromatic rings of adjacent dimers, as shown in Figure 4.9.[1] The copper-halogen

[1]An ideal model would also include the hydrogen bonding interactions included in the previous model however, to maintain charge neutrality, this would require the addition of two extra cations to

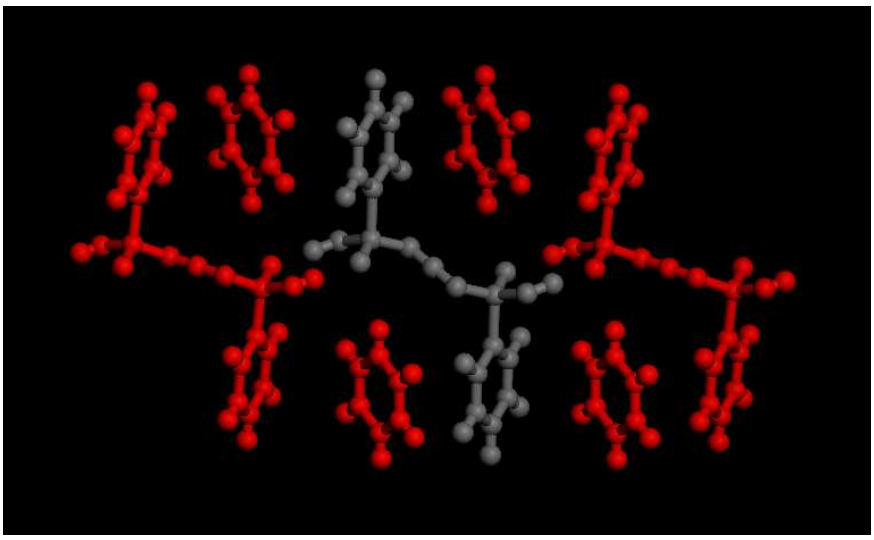

Figure 4.8: Unsubstituted single H-bond phenylphosphonic acid dimer with aromatic rings from adjacent dimer molecules and hydrogen bonding dimers. Red atoms have been constrained in their experimental positions while grey atoms have been allowed to relax, without constraints.

cation and aromatic rings have been constrained while the target dimer was allowed to relax. In contrast to the previous two models, the O $\cdots$ O distance increases to 2.475 Å. The C-P-O-H angle also increases to 94.119°. It is tempting to conclude that electrostatic forces induced by the copper-halogen cation act to lengthen the O $\cdots$ O distance. However, the model shown in Figure 4.8 also contains electrostatic interactions in the form of hydrogen bonding anions. Hence, it is more likely that the presence of an unpaired electron (associated with the Cu^{2+} cation) in the current model is responsible for the observed lengthening. This hypothesis is supported by masNMR data (unpublished), which show an interaction between the phosphorus atoms of the anion and the copper centre.

The C-P-O-H torsion angle appears to be most sensitive to crystal field effects, in agreement with Kitaigorodsky's hypothesis. In a recent study [122] of the effects of the crystal field on crystal structures containing metal complexes, torsion angles showed deviations of up to 47°. The O $\cdots$ O bond length also appears to be affected by the crystal environment, in apparent contradiction of Kitaigorodsky's hypothesis but in support of the findings of recent studies. [122, 123] The discrepancies between experi-

the current model, *i.e.* three phenylphosphonic acid anions balanced by three copper-halogen cations. Calculations on such a model were attempted but proved too computationally complex.

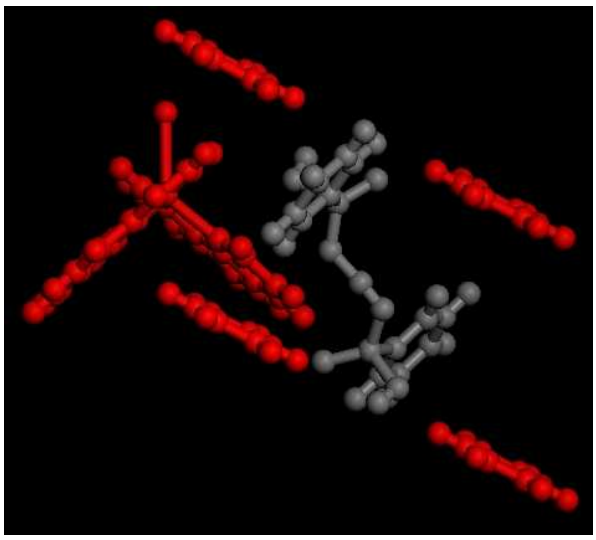

Figure 4.9: Unsubstituted single H-bond phenylphosphonic acid dimer with aromatic rings from adjacent dimer molecules and copper-halogen cation. Red atoms have been constrained in their experimental positions while grey atoms have been allowed to relax, without constraints.

ment and theory in each model considered were less than 0.01 Å however, this is still larger than the uncertainty in the x-ray structure (by about one order of magnitude). Overall, the results of this section suggest that the observed experimental geometry of the single H-bond dimer is a complex product of competing intra- and intermolecular forces and crystal field effects.

4.4.4 Cooperative Phenomena

The formation of a hydrogen bond induces changes in both the geometric and electronic structures of the participating sub-units. For example, it is often observed that a chain of hydrogen-bonded molecules is more strongly bound than an individual hydrogen-bonded dimer would be in isolation. This effect is but one example of the 'cooperative' nature of hydrogen bonds and it is this cooperativity that leads to the common occurrence of assemblies of hydrogen-bonded molecules in solids. [124]

In order to investigate cooperative effects in the organophosphonate family,[2] we first constructed a toy model consisting of a chain of six pairs of copper-halogen cations and phenylphosphonic acid dimer anions, as shown in Figure 4.10 for the iodo-form. The

[2]Recall that the four members of the organophosphonate family are distinguished by the halogen bound to the copper ion. The crystals are structurally identical in every other respect.

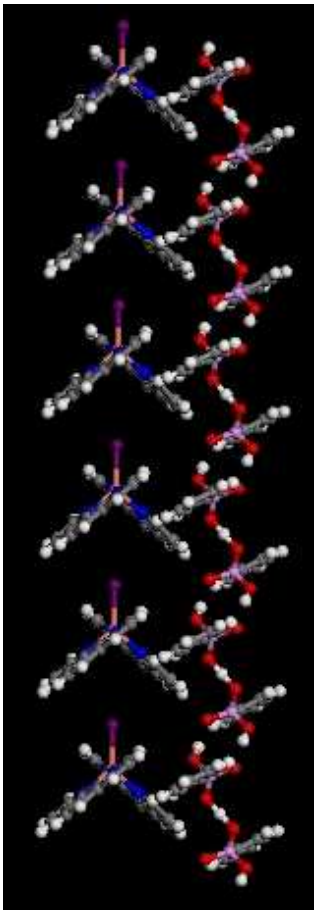

Figure 4.10: A fragment of six cation-anion pairs along the b axis for the iodo-form.

chain is a one-dimensional fragment of the b axis of the crystal; all atoms are in their experimental positions. An unpaired spin was associated with each cation-anion pair (the copper ion is in a +2 oxidation state and is therefore paramagnetic) although the entire model is electrically neutral. To reduce the computational effort required, only single-point energies were calculated.

A meaningful interpretation of the results requires that we define the term 'interaction energy' for the systems being studied in this work. Figure 4.11 shows two cation-anion pairs for the iodo-form, where the dashed green line denotes the 'interaction interface'. We can see that the hydrogen bonds linking the phenylphosphonic acid dimers, electrostatic interactions between the copper complex cations and phenylphosphonic acid dimer anions and dispersion forces will all contribute to the interaction energy to vary-

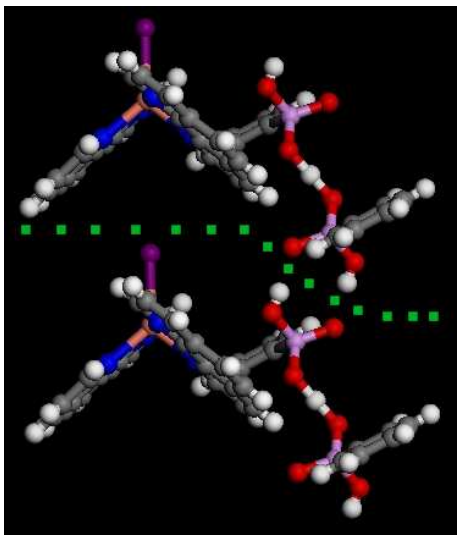

Figure 4.11: A cation-anion pair for the iodo-form. The dashed green line denotes the 'interaction interface'.

ing degrees. Our aim in this section is to understand how the addition of cation-anion pairs to a chain affects interactions between each cation-anion pair in the chain at the interaction interface.

We have calculated one-body, two-body and three-body interaction energies, which we have defined as follows.
The one-body interaction energy for n cation-anion pairs is defined as,

$$E^1_{int}(n) = [E^n_{chain} - n(E_{pair})] - E^1_{int}(n-1), \tag{4.2}$$

where E^1_{int} is the one-body interaction energy, E^n_{chain} is the total energy of a chain of n cation-anion pairs and E_{pair} is the total energy of a single cation-anion pair. For $n = 2$, $E^1_{int}(n-1) = 0$ and (4.2) simply reduces to the expression in square brackets. For $n \geq 3$, (4.2) is calculated iteratively. We have calculated $E^1_{int}(n)$ for $n = 3 - 6$.
We have also calculated two-body and three-body interaction energies using expressions similar to (4.2). For the sake of clarity, we have explicitly defined n in the following equations. For two cation-anion pairs, the two-body interaction energy is defined as,

$$E^2_{int}(4) = E^4_{chain} - 2E^2_{chain}, \tag{4.3}$$

Table 4.7: Many-body energies (in kcal) as a function of number of units for each member of the organophosphonate family. A unit is defined as a cation-anion pair.

Units	**I**	**Cl**	**Br**	**NCS**
ONE-BODY				
2	-12.29	-7.47	-15.42	-19.38
3	-11.72	-6.85	-14.78	-18.77
4	-11.53	-6.66	-14.59	-18.60
5	-11.46	-6.57	-14.50	-18.52
6	-11.41	-6.52	-14.46	-18.48
TWO-BODY				
4	-10.96	-6.04	-13.95	-17.98
6	-10.57	-5.62	-13.55	-17.62
THREE-BODY				
6	-10.39	-5.43	-13.36	-17.44

where $E^2_{int}(4)$ is the two-body interaction energy for 4 cation-anion pairs, E^4_{chain} is the total energy of a chain of 4 cation-anion pairs and E^2_{chain} is the total energy for a chain of two cation-anion pairs.

For three cation-anion pairs, the two-body interaction energy is defined as,

$$E^2_{int}(6) = (E^6_{chain} - 3E^2_{chain}) - E^2_{int}(4). \tag{4.4}$$

Finally, the three-body interaction energy is defined as,

$$E^3_{int}(6) = E^6_{chain} - 2E^3_{chain}. \tag{4.5}$$

Results are shown in Table 4.7.

An immediate observation from Table 4.7 is that the interaction energies vary according to the halogen attached to the copper ion, with NCS having the strongest interaction energies and Cl the weakest. There does not appear to be any correlation between the magnitude of the interaction energy and the size of the halogen. It is also reasonable to infer that dispersion is not the main contribution to the interaction energy. If this were the case, we would expect the iodo-form to have the largest interaction energy, since iodine is the largest of the halogens considered and therefore the most polarisable.

The interaction energies are essentially converged by the addition of the third unit

Table 4.8: Temperature range, in degrees Celsius, in which each member of the organophosphonate family begins to decompose. citeclarke02

I	Cl	Br	NCS
220-350^o	40-350^o	230-380^o	250-380^o

for the one-body interactions and vary only little for the two-body and three-body interactions. This suggests that the interaction between cation-anion pairs is mostly local and one-body in nature.

Contrary to our discussion of cooperative effects at the beginning of this section, the interaction energy *decreases*, rather than increases, as the chain grows, that is, the chain becomes less stable with the addition of cation-anion pairs. This sort of weakening is commonly referred to as negative cooperativity and arises when a molecule acts as a double proton donor and proton acceptor. Inspection of Figure 4.11 shows that the phenylphopshonic acid molecules bear two OH groups each and this allows them to act as double proton donors and acceptors. The slight destabilisation caused by such an arrangement must be outweighed by stabilising forces not included in the current model.

It would be desirable to relate the calculated interaction energies with some observable physical property of the organophosphonate materials. The results of thermogravimetric analysis (TGA) experiments [4] may be helpful in this respect. In a TGA experiment, the mass of a sample in a controlled environment is recorded as a function of temperature versus time. The stepwise mass losses can be used to calculate sample decomposition reactions of volatile, organic and inorganic constituents. Table 4.8 shows that there appears to be a reasonable correlation between the onset of melting and the calculated interaction energies. Cl was calculated to have the weakest interaction energy and has the lowest melting-onset temperature.[3] Conversely, NCS was calculated to have the strongest interaction energy and has the highest melting-onset temperature. While the experimental results are certainly not definitive, they do serve as a useful guide.

[3]We note that, in the case of Cl, melting temperatures of less than 100^{o}C usually indicate the presence of surface, non-chemically bound water. Having taken this fact into account, we have assumed (given the results for Br, I and NCS) that the melting-onset temperature will still be lowest for Cl.

4.5 Conclusions

In this chapter, we calculated the binding energies and geometries of substituted and unsubstituted phenylphosphonic acid dimers. We also considered the effect of various parts of the crystal field on the geometry of a single, unsubstituted phenylphosphonic acid dimer anion and analysed cooperative phenomena as a function of chain length for each member of the organophosphonate family.

The optimised geometries of each dimer system studied were in good agreement with the experimentally determined geometries. The calculated binding energies were consistent with those of other ionic hydrogen bonded systems. The binding energy of the methyl-substituted dimer was close to that of its unsubstituted parent, as expected given the similarity between the structures of the two dimers and their respective electron density distributions. The nitro-substituted dimer was found to have a lower binding energy than both the methyl and unsubstituted dimers. The crystal packing arrangements for each of the three different dimers are similar. We are thus tempted to conclude that the structure is robust to changes in the phenylphosphonic acid. However, it is unclear whether differences in crystal packing are caused by the the identity of the substituent or merely its position on the aromatic ring (ortho-, meta- or para-). The work of Nagarajan, *et al* [120] appears to suggest that the identity of the substituent is not as important as its position on the ring. Further studies are currently underway to address this issue.

The C-P-O-H torsion angle was found to be the most sensitive geometric parameter to crystal field effects, in agreement with Kitaigorodsky's hypothesis. Different parts of the crystal field were found to affect the O$\cdots$O bond length in different ways. When the dimer was simulated in isolation, the O$\cdots$O distance contracted and this did not change with the inclusion of hydrogen bonding interactions. In contrast, the inclusion of electrostatic effects (in the form of a copper complex cation) increased the O$\cdots$O distance. Larger, more sophisticated models may shed more light on this important problem.

Finally, we found that interaction energies between cation-anion pairs varied according

to the halogen attached to the copper ion: NCS was found to have the strongest interaction energies and Cl the weakest. One feature common to each member of the organophosphonate family is that the interaction energies are converged by the addition of the third unit for the one-body interactions, with little variation between the two-body and three-body interactions. This ultimately suggests that interactions between cation-anion pairs are mostly one-body and local in nature.

Chapter 5

Metal-Ligand Interactions in the Copper-Halogen Cations

5.1 Introduction

In the previous chapter we looked in detail at the intermolecular hydrogen bonding interaction between the phenylphosphonic acid molecules in the organophosphonates. In this chapter we turn our attention to the intramolecular metal-ligand interactions that characterise the other component of the organophosphonates under study in this thesis – the copper-halogen-phenanthroline cations.

As mentioned in Chapter 1, the iodo-form of the organophosphonate family displayed some curious properties, which were different to those of the other complexes. In addition to some initial uncertainty concerning the charge of the copper ion, the iodo-form was the only species for which the starting material could be isolated and it was found to exist in two polymorphic forms [2]; one of these polymorphs is shown in Figure 5.2. Both types of polymorph contain ribbons of $[Cu(C_{12}H_8N_2)_2]^+$ cations together with unassociated I^- positioned in the voids, *i.e.* the iodine is not bound to the copper ion in either polymorph and the copper ion exists in a +1 oxidation state. The reaction of these complexes with phenylphosphonic acid in a hydrothermal

synthesis at 160°C produced the structures under study in this work, [3] as shown in Figure 5.1.

$$[\mathrm{Cu(phen)_2}]^{+}\cdot\ \mathrm{I}^{-} \xrightarrow{\mathrm{A,B}} [\mathrm{CuI(phen)_2}][(\mathrm{C_6H_5PO(OH)_2})(\mathrm{C_6H_5PO_2(OH)}]$$

Figure 5.1: Reaction scheme showing conversion of Cu(I) starting material to Cu(II) product, where A=heat, B=phenylphosphonic acid and phen=1,10-phenanthroline.

During the course of the reaction, the copper ion undergoes a change of oxidation state, from +1 to +2. The I^- counter-ion simultaneously binds to the copper ion, thus forming a five- instead of four-coordinate complex.[1] The theoretical simulation of this reaction will form one aspect of this chapter.

That there are structural differences between Cu(I) and Cu(II) complexes is well known [125] and is an interesting property inherent to copper coordination compounds. Cu(I) has a d^{10} electronic configuration and is typically four-coordinate with approximate tetrahedral geometry. Cu(II) is d^9 and adopts a Jahn-Teller distorted geometry that is usually 5- or 6-coordinate. The structures of the Cu(I) polymorph starting material and the Cu(II) organophosphonate product are consistent with these observations.

An understanding of the stereochemical preferences of Cu(I) and Cu(II) compounds is relevant to their use in photochemical applications. Cuprous bis-phenanthroline compounds possess metal-to-ligand charge transfer (MLCT) excited states, which make them attractive candidates for photochemical devices, such as solar energy converters, chemical sensors and photocatalysts. Complexes of ruthenium(II), osmium(II) and rhenium(I) have traditionally been employed for these applications however, these elements are generally expensive and environmentally hazardous. In contrast, copper is less toxic and less expensive and thus represents a viable alternative to ruthenium-based devices.

The majority of copper(I) photochemical research has focused on the complexes of disubstituted phenanthrolines because the MLCT states in these systems are long-

[1] It is interesting to note that although two polymorphs were identified for the starting material, only one product was isolated.

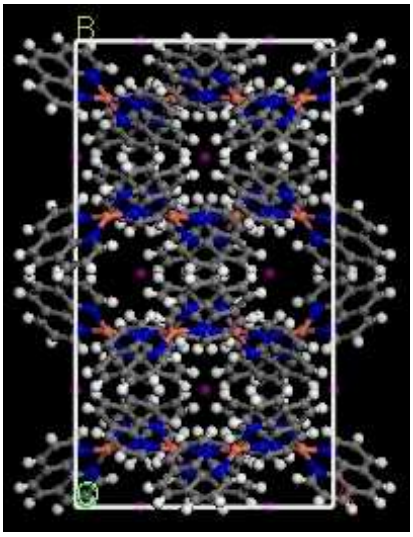

Figure 5.2: A single unit cell of a Cu(I)bis-(1,10-phenanthroline) polymorph. Different atoms are distinguished by the following colours: grey–carbon, white–hydrogen, blue–nitrogen, orange–copper, purple–iodine.

lived, even at room temperature. In contrast, the parent (unsubstituted) compound is non-emissive under the same conditions and has therefore been less well studied. Although we have not considered the excited state properties of the Cu(I) and Cu(II) complexes, knowledge of the stereochemical preferences of Cu complexes in different oxidation states is still relevant to our original aim of using theoretical methods to understand interactions in complex environments.

5.2 Previous Studies

Complexes of transition metals, such as the copper complexes described above, are often characterised by charge-transfer interactions and have unpaired electrons. The theoretical treatment of such systems, particularly within DFT, is complicated and fraught with problems. In this section, we will discuss the main problems in the context of previous studies.

The modern theory of charge-transfer complexes was pioneered by Mulliken. [126] In his approach, the ground state wavefunction of a charge-transfer complex is written as

$$\Psi_{DA} = a\,\psi_0(D, A) + b\,\psi_1(D^+, A^-), \tag{5.1}$$

where $a \gg b$ and D and A refer to the electron donor and acceptor, respectively. In physical terms, equation (5.1) implies that the charge-transfer complex may be considered a mixture of an ionic pair state, $\psi_1(D^+, A^-)$, and a non-ionic pair, $\psi_0(D, A)$. Mulliken's theory successfully explained many spectroscopic and structural results however, it failed to predict the stability of some complexes. It was later argued that the charge-transfer interaction alone was not sufficient to describe all properties of charge-transfer complexes. Subsequent energy decomposition studies by Morokuma [127] showed that interactions other than simple charge-transfer, particularly electrostatic, are important.

The primary charge-transfer interaction in the copper-iodine-phenanthroline cations is the Cu-I bond.[2] The unique and interesting electronic structure of the transition metal dihalides, particularly the copper dihalides, has made them the subject of extensive theoretical and experimental studies. Wang and Schwarz [128] undertook a density functional study of the first row transition metal dihalides with the dual aim of studying the electronic properties of the ground states and assessing the performance of DFT for these types of systems. They found that DFT performed adequately for structural and vibrational properties but consistently over-estimated dissociation energies (too-strong binding) by about 0.5-1 eV. This is consistent with an earlier DFT study of ethylene-halogen complexes by Ruiz, *et al.* [129] The authors found that interaction energies were strongly overestimated within both the LDA and GGA. They attributed this behaviour to the too-rapid asymptotic decay of the exchange-correlation potential, which is associated with an incomplete cancellation of the self-interaction error.

In Section 5.4.2 of this chapter, we consider the dissociation of the charge-transfer complex $[CuI]^+$. The self-interaction error can be particularly pronounced in dissociating ('stretched') odd-electron systems, leading as before to an over-estimation of the dissociation energy. In addition, the dissociation curves for such systems are commonly observed to contain spurious energy barriers. Bally and Sastry [130] calculated the dissociation curves for $H_2{}^+$ and $He_2{}^+$ using UCCSD(T), UHF and a number of different exchange-correlation functionals. The DFT methods predicted an unphysical energy barrier at large inter-fragment distances and tended asymptotically to energies

[2] Charge-transfer between the copper ion and phenanthroline ligands may also be important.

lower than that of the bound species, *i.e.* the separated fragments were predicted to be lower in energy than the bound dimers. Along with other authors [131], they also noted that with the DFT methods, the spin and charge distributions of the dissociating species remained completely uniform (delocalised) at all inter-fragment distances. The origin of this problem is commonly attributed to the the incorrect description of the exchange-correlation hole in stretched systems. All current approximate functionals assume that the exchange-correlation hole is localised. However, in molecular systems, particularly stretched molecular systems, the exchange hole may be delocalised over several centres. Inherently local approximate functionals cannot reproduce this behaviour. Though various remedies have been proposed [63, 132, 133, 134], the self-interaction and exchange-correlation hole errors still plague systems with unpaired electrons.

The list of problems discussed thus far is by no means complete however, it is beyond the scope of this work to consider the matter further here. We will consider additional problems as they arise throughout the chapter and refer the reader to the following reference for a more comprehensive discussion. [96]

5.3 Computational Details

5.3.1 System, Functional and Basis Set Selection

We have taken as our model systems both the starting material Cu(I) polymorph and the Cu(II) organophosphonate product (iodo-form) cations, as shown in Figure 5.3. The initial co-ordinates for each structure were obtained from their respective bulk crystals, the structures of which were determined by single-crystal x-ray diffraction. [2, 3] The geometries of each model were optimised, without constraints, using the PBE functional and DNP basis with scalar relativistic corrections. Each cation was simulated in isolation, *i.e.* we have neglected the effects of the crystal environment and solvent. All calculations were spin-unrestricted.

In Chapter 1 we described how each member of the organophosphonate family contains a Cu(II) cation co-ordinated to two phenanthroline ligands and a halogen substituent, either I^-, Br^-, Cl^- or NCS^-. Thus far we have mainly focussed on the iodo-form.

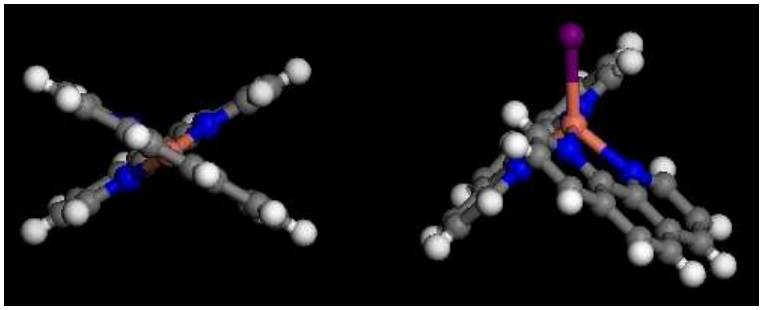

Figure 5.3: Copper cation model systems. Left: Cu(I) starting material. Right: Cu(II) product.

In Section 5.4.3, we analyse and compare the charge-transfer interaction of the iodo-, chloro- and bromo-forms. We have performed an analysis of the charge on the Cu and halogen sites using the Mulliken scheme. We note that, since there is no quantum mechanical operator for atomic charge, any charge partitioning scheme will be necessarily arbitrary. In the case of the Mulliken analysis, the results are also dependent on the basis set used. In the current calculations the use of a consistent basis set means that variations in the population analysis with halogen do provide a useful guide to changes in the nature of the electronic state. The initial co-ordinates for each structure were obtained from the respective experimental bulk crystals and the geometries of each model were optimised using the PBE functional and DNP basis (with scalar relativistic corrections, spin-unrestricted). Each cation was simulated in isolation and Mulliken charges calculated for the relaxed structures.

5.4 Results and Discussion

5.4.1 Relaxations and Electronic Structure – Cu(I) Starting Material and Cu(II) Product Cations

The isolation of the Cu(I) starting material provides us with a unique opportunity to compare the geometry and electronic structure of two very similar molecules in two different crystal environments. Inspection of Figure 5.3 reveals that the geometries of each cation are markedly different, consistent with previous observations for Cu(I) and Cu(II) complexes. To what extent is the difference attributable to the electronic structure of the Cu ion or the crystal environment?

Table 5.1: Comparison of main geometric parameters between experimental [2] and optimised (PBE/DNP) structures for Cu(I) starting material.

	Experiment	**Optimised**
d(Cu–N1), Å	2.038	2.028
d(Cu–N2), Å	2.032	2.028
d(Cu–N3), Å	2.047	2.026
d(Cu–N4), Å	2.037	2.024
N1–Cu–N2,°	114.435	123.557
N2–Cu–N3,°	82.332	82.812
N3–Cu–N4,°	114.455	126.540
N4–Cu–N1,°	82.273	82.738

Table 5.2: Comparison of main geometric parameters between experimental [3] and optimised (PBE/DNP) structures for Cu(II) product.

	Experiment	**Optimised**
d(Cu–N1), Å	2.123	2.179
d(Cu–N2), Å	1.983	2.039
d(Cu–N3), Å	2.123	2.179
d(Cu–N4), Å	1.983	2.039
d(Cu–I), Å	2.625	2.641
N1–Cu–N2,°	95.238	95.356
N2–Cu–N3,°	80.408	79.405
N3–Cu–N4,°	95.238	95.356
N4–Cu–N1,°	80.408	79.405

We begin by comparing the geometry of the optimised structures with that of the experimental ones for both the Cu(I) starting material and Cu(II) product, as shown in Tables 5.1 and 5.2. In the case of the Cu(II) product, the agreement between the optimised and experimental structures is excellent. Theory correctly predicts a structure with C2 symmetry.

In the case of the Cu(I) starting material, the agreement is good as far as bond lengths are concerned however, the agreement for bond angles is less good. Although theory correctly predicts a structure with C1 symmetry, the optimised structure is closer to ideal tetrahedral geometry than the experimental one, *i.e.* the experimental structure is distorted from ideal tetrahedral geometry, consistent with the observations of several authors. [135, 136, 137] This discrepancy could be attributed to crystal packing effects, which are absent from the model currently under study. This hypothesis is supported by proton nuclear magnetic resonance data of these complexes in solution [138], which indicate equivalence of the two ligands at room temperature. The optimised structure, with two equivalent phenanthroline ligands, is consistent with these observations.

In this instance, it is easier and less computationally intensive to simulate changes in electronic structure than crystal field effects. Hence, for the next step in our investigation, we constructed a new model system consisting of a single Cu cation in the experimentally determined product geometry. The iodine was removed and the electronic configuration at the Cu ion was changed from that of the product, d^9, to that of the starting material, d^{10}. The system charge was set to +1. Hence, the model has the geometry of the product Cu(II) ion but the electronic configuration of the Cu(I) starting material. The geometry was optimised without constraints; results are shown in Table 5.3. Note that we take as our reference for comparison the optimised structure of the Cu(I) starting material, not the experimentally determined structure. This is to keep all comparisons consistent, *i.e.* all our simulations are being carried out in the gas-phase while the experimentally determined structures come from solid-state crystal structure data.

Table 5.3 shows that the agreement between the Model and Reference geometric parameters is excellent. One discrepancy is that the Model has C2 symmetry while the

Table 5.3: Comparison of main geometric parameters between cation with product geometry and starting material electronic configuration ('Model') and optimised Cu(I) starting material ('Reference').

	Model	**Reference**
d(Cu–N1), Å	2.017	2.028
d(Cu–N2), Å	2.031	2.028
d(Cu–N3), Å	2.031	2.026
d(Cu–N4), Å	2.017	2.024
N1–Cu–N2,°	122.720	123.557
N2–Cu–N3,°	82.741	82.812
N3–Cu–N4,°	122.720	126.540
N4–Cu–N1,°	82.741	82.738

Reference has C1 symmetry. This is because the Cu(II) product cation has C2 symmetry while the Cu(I) starting material has C1 symmetry. The optimised Model has retained its initial symmetry. The results show that we can make the product cation 'look like' the starting material by changing the electronic configuration of the Cu ion to that of the starting material.

A logical next step would be to see whether the above experiment can work in reverse, *i.e.* can we make the starting material look like the product by changing the electronic configuration of the Cu ion? To answer this question, we constructed another model system consisting of a single Cu cation in the experimentally determined starting material geometry. An iodine atom was bound to the copper ion at a distance of 2.6 Å, approximately the Cu-I length found for the experimental structure. The electronic configuration of the Cu ion was set to d^9 and the system charge set to +1 (the iodine carries a charge of -1). This model has the geometry of the starting material (with an added iodine atom) but the electronic configuration of the Cu(II) product. Here as before, the geometry was optimised, without constraints. Results are shown in Table 5.4. Again we take as our reference for comparison the optimised structure of the Cu(II) product, not the experimentally determined structure.

Table 5.4 shows that the agreement between the Model and Reference is again very good. As before, the symmetry of the starting material (C1) has been retained in the optimised Model. Hence, we can make the starting material look like the product by altering the electronic configuration of the Cu ion. It appears that the answer to

the question posed at the beginning of this section is that the difference in geometry between the Cu(I) starting material and Cu(II) product cation is largely attributable to the electronic configuration of the Cu ion, not the crystal environment.

5.4.2 Charge-Spin Separation in the Dissociating $[CuI]^+$ System

In the previous section we studied the discrete end-points of the Cu(I)→Cu(II) reaction. In this section, we attempt to monitor changes in the electronic structure of the system as the iodine moves away from the copper ion.[3] To simplify the problem the phenanthroline rings were removed from the system to leave $[CuI]^+$. We have thus eliminated the need for geometry optimisation at each step and only consider charge-transfer between the copper ion and iodine atom (this is a justifiable assumption, as we shall see in Section 5.4.3). Hence, the reaction we are simulating is $Cu^{2+}I^- \rightarrow Cu^+ + I$.

The dissociation of ionic systems with unpaired electrons requires the separation of charge and spin. In theory, as the two sub-systems move further apart, an electron is transferred from one fragment to the other. At large distances between the fragments, the positive charge should concentrate at the ionic fragment (Cu^+ d^{10}, in our case) and the spin density at the radicalic one (I p^5). This phenomenon is known as charge-spin separation. In Chapter 1 we noted that the the dissociation of odd-electron systems presents one of the biggest challenges to practitioners of DFT. In our case, the presence of a transition metal further complicates the problem.

[3] In reality, the iodine becomes bound to the copper during the course of the reaction however, for technical reasons, dissociation is easier to simulate than association.

Table 5.4: Comparison of main geometric parameters between cation with starting material geometry and product electronic configuration ('Model') and optimised Cu(II) product ('Reference').

	Model	**Reference**
d(Cu–N1), Å	2.185	2.179
d(Cu–N2), Å	2.028	2.039
d(Cu–N3), Å	2.172	2.179
d(Cu–N4), Å	2.027	2.039
N1–Cu–N2,°	97.991	98.356
N2–Cu–N3,°	79.618	79.405
N3–Cu–N4,°	97.430	98.356
N4–Cu–N1,°	79.702	79.405

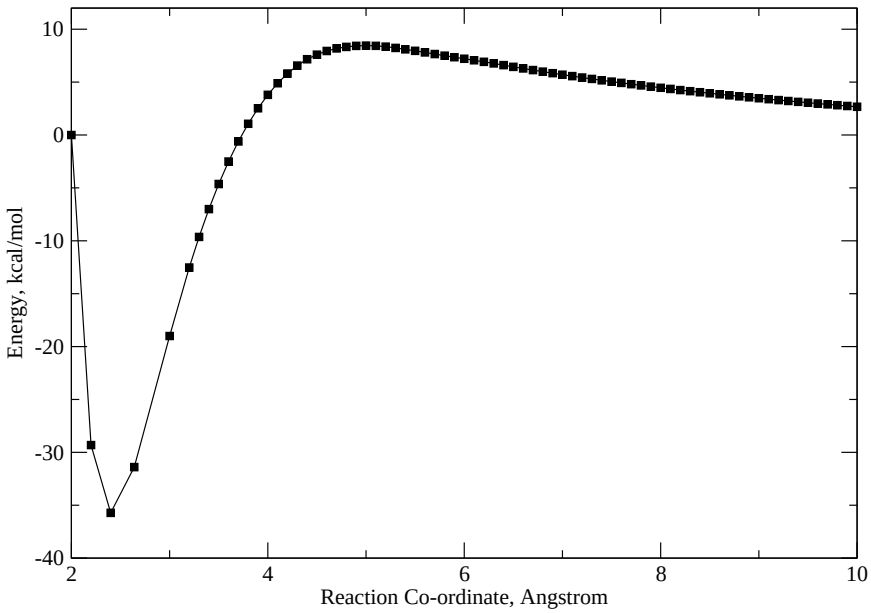

Figure 5.4: Dissociation curve for the reaction $Cu^{2+}I^{-} \rightarrow Cu^{+} + I$.

The Cu–I bond length was taken as the reaction co-ordinate for the dissociation reaction. The two atoms were moved apart in 0.05 Å increments and the total energy calculated at each step using the PBE functional and DNP basis with scalar relativistic corrections. All calculations were spin-unrestricted. The dissociation curve obtained is shown in Figure 5.4.

Consistent with the observations of other authors, Figure 5.4 features a fictitious transition state at $\sim$ 4.5 Å and there is a subsequent decrease in energy from 5-10 Å. As noted earlier, this unphysical behaviour is commonly observed for dissociating odd-electron systems and is accompanied by a severe over-estimation of the dissociation energy. The origin of this error is a complex mixture of inter-related problems, which have been much studied and debated in the literature. The unphysical shape of the dissociation curve is thought to be due to the incorrect description of the exchange-correlation hole in approximate functionals [132, 139, 133, 134], while the over-estimated dissociation energy has been attributed to self-interaction error. [140, 133, 134] The self-interaction error is present in all systems with unpaired electrons but becomes negligible in large, homogeneous systems. The $[CuI]^{+}$ system contains reasonably heavy atoms and hence the error is not as pronounced as for H_2^+ and He_2^+. Bally and Sastry [130] showed

that for the hydrogen and helium dimer cations, some exchange-correlation functionals (most notably, BLYP) produced dissociation curves which showed that the two separated sub-systems were more stable than the bound dimers. Functionals which contain a fraction of exact exchange ('hybrids') fail less badly because they are better able to describe delocalised exchange holes. Calculations with a hybrid functional were not attempted for $[CuI]^+$ because only 'pure' functionals are supported in DMol^3; the use of other codes to perform the calculations was investigated however, we were unable to find a suitable basis set and pseudopotential for iodine.

Another difficulty encountered during the course of these simulations, not evident from Figure 5.4, is that the HOMO-LUMO gap decreased with increasing separation, which subsequently lead to severe SCF convergence difficulties. Such behaviour may indicate the presence of a state with multireference character [141] and is often observed in systems where the bonds are stretched. In such situations, the ground state is not a purely covalent state (for example) but is a mixture of covalent, ionic and possibly higher states. The accurate description of such states requires the use of multiple determinants, which significantly increases the cost of the calculation. In our case, a smearing technique [142] was applied to aid convergence however, some authors [96] have questioned the physical validity of such solutions.

Are any currently available methods able to overcome these problems? The Perdew-Zunger self-interaction correction scheme does, in theory, solve the self-interaction problem. However, it is computationally intensive for large or heavy systems and remains largely untested for systems other than atoms and chemically simple molecules. The high-level Coupled Cluster technique does not suffer from self-interaction error but is also impractical for systems with more than approximately ten electrons. The use of pseudopotentials to replace core electrons can assist in making problems computationally tractable however, in our case, the replacement of the iodine core electrons would still leave a large number of electrons. The Quantum Monte Carlo technique is also free from self-interaction error and can be used to study much larger systems than are possible with Coupled Cluster. The problem here is that the final accuracy attainable in a QMC calculation depends crucially on the quality of the trial wavefunction, which

is usually obtained from a Hartree-Fock or DFT calculation. If the trial wavefunction is self-interaction contaminated, it is unlikely to be of sufficient quality to produce meaningful results in a subsequent QMC calculation. Hence, although this investigation was not successful, the $[CuI]^+$ system demonstrates some of the many challenges involved in simulating open-shell systems, particularly dissociating open-shell systems.

5.4.3 Charge-Transfer Interactions in the Organophosphonate Family

So far we have only considered charge-transfer interactions in the iodo-form. In this section, we place the iodo-form results in context by analysing and comparing the charge-transfer interaction in the chloro- and bromo-forms. Each member of the organophosphonate family is structurally identical, except for the halogen attached to the copper ion. Hence, any change in electronic structure between the three complexes is solely attributable to a change in halogen.

The charge-transfer process is usually mediated by interactions between the frontier orbitals of the participating donor and acceptor species. Since charge-transfer is essentially an electron exchange process, the geometry of the complex must be such that there is maximal overlap of the HOMO and LUMO. Hence, for a pure charge-transfer interaction, the interaction energy would be directly proportional to the energy difference between the donor and acceptor orbitals.

Halogens can function as both electron-acceptors and electron-donors. The vacant antibonding σ_u orbital of the halogen is available to accept electrons while the occupied π_g orbitals afford a π-donor system. Experiments have established that the strength of the charge-transfer interaction for halogen molecules increases as one goes from fluorine to iodine, observations which have been supported by results from theoretical calculations. In their study of ethylene-halogen complexes, Ruiz *et al* [129] attributed this behaviour to the energetic properties of the LUMO of the acceptor halogen molecules. Wang and Schwarz [128] noted that for the transition metal dihalides, the π-donating property of the halogen ligands destabilises the $3d\pi$-type orbital in the molecules. If they become occupied by electrons, the Cu–X bond lengthens.

As before, we begin by comparing the geometries of the optimised structures with that of the experimentally determined ones. The initial co-ordinates for each structure were obtained from the respective bulk crystals and the geometry optimised using the PBE functional and DNP basis. All calculations were spin unrestricted. The geometry of the optimised iodo-form appears in Table 5.2 however, we reproduce it in Table 5.5 to make comparison with the chloro- and bromo-forms easier.

Table 5.5 shows that the agreement between experiment and theory is generally excellent. Cu–N bond lengths appear to slightly decrease with increasing halogen size, while Cu–X bond lengths increase with increasing halogen size. Each complex has C2 symmetry.

The increasing Cu–X bond length on going from chlorine to iodine has two related effects at its origin. Charge-transfer tends to lengthen bonds and the π-donating ability of the halogens increases going down a period. Iodine is the best π-donor; the Cu–I bond is therefore the longest. A decrease in the HOMO-LUMO gap with increasing π-donating ability is another manifestation of this phenomenon,[4] as shown in Table 5.6, which also contains the computed Mulliken charges for each species. Table 5.6 shows that as the amount of (negative) charge transferred onto the Cu centre increases, the HOMO-LUMO gap decreases.[5]

Finally, we present experimental data in support of our theoretical calculations. The solid-state UV-visible spectra of the organophosphonates have been recorded. [4] The spectra show two bands: band A is attributed to phenanthroline absorption and occurs at approximately the same frequency for each complex; band B, which varies from chlorine through to iodine, is attributed to halide-to-metal charge-transfer. For a series of similar halide complexes of a given metal ion, it has been found that the charge-transfer absorption shifts to lower frequencies as the π-donating ability of the halide increases from Cl^- to I^-. This is observed in the complexes under study in this thesis.

[4] In the language of crystal field theory, good π-donors decrease the crystal field splitting parameter, which is analogous to the HOMO-LUMO gap.

[5] We note that the DFT band gap problem, which results in band gaps being underestimated for insulators and semi-conductors, affects HOMO-LUMO gaps in the same way in molecular calculations. This should not concern us too much, since we are only interested in qualitative trends.

Table 5.5: Comparison of main geometric parameters between experimental and optimised (PBE/DNP) structures for chloro-, bromo- and iodo-forms.

	Experiment	**Optimised**
X=Cl		
d(Cu–N1), Å	2.146	2.195
d(Cu–N2), Å	1.991	2.026
d(Cu–N3), Å	2.146	2.195
d(Cu–N4), Å	1.991	2.026
d(Cu–X), Å	2.260	2.246
N1–Cu–N2,°	95.995	98.552
N2–Cu–N3,°	80.064	79.134
N3–Cu–N4,°	95.995	98.552
N4–Cu–N1,°	80.064	79.134
X=Br		
d(Cu–N1), Å	2.137	2.190
d(Cu–N2), Å	1.988	2.023
d(Cu–N3), Å	2.137	2.190
d(Cu–N4), Å	1.988	2.023
d(Cu–X), Å	2.421	2.407
N1–Cu–N2,°	95.680	98.393
N2–Cu–N3,°	80.051	79.440
N3–Cu–N4,°	95.680	98.393
N4–Cu–N1,°	80.051	79.440
X=I		
d(Cu–N1), Å	2.123	2.179
d(Cu–N2), Å	1.983	2.039
d(Cu–N3), Å	2.123	2.179
d(Cu–N4), Å	1.983	2.039
d(Cu–X), Å	2.625	2.641
N1–Cu–N2,°	95.238	95.356
N2–Cu–N3,°	80.408	79.405
N3–Cu–N4,°	95.238	95.356
N4–Cu–N1,°	80.408	79.405

Table 5.6: Comparison of the HOMO-LUMO (H-L) gap and charge state of the Cu and X atoms computed using a Mulliken scheme.

	Charge	**H-L gap, eV**
Cu	+0.501	-1.006
Cl	-0.422	
Cu	+0.496	-0.835
Br	-0.412	
Cu	+0.431	-0.600
I	-0.363	

Table 5.7: UV-visible data (10^{-3}cm^{-1}) for the copper complexes [4].

Complex	Band A	Band B
Cu–Cl	39.21	27.24
Cu–Br	39.37	23.86
Cu–I	39.52	21.73

UV-visible data are presented in Table 5.7.

5.5 Conclusions

In this chapter, we investigated the stereochemical preferences of the Cu(I) and Cu(II)-phenanthroline cations. We also attempted to monitor changes in the electronic structure of the $[CuI]^+$ system by simulating the dissociation reaction $Cu^{2+}I^- \rightarrow Cu^+ + I$. In the final section of this chapter, we compared the charge-transfer interaction for the chloro-, bromo- and iodo-forms of the organophosphonate system.

The optimised geometries of each model considered were in good agreement with the experimentally determined geometries. The discrepancy between theory and experiment for the Cu(I) model was attributed to crystal packing effects, which were absent from the model being studied. We also found that the geometry of the Cu complex could be altered by changing the electronic configuration of the Cu ion. This suggests that the geometric differences between the Cu(I) and Cu(II) complexes are mostly due to changes in the electronic configuration of the Cu ion, not crystal packing effects.

We were unsuccessful in our attempt to simulate the dissociation reaction $Cu^{2+}I^- \rightarrow Cu^+ + I$, although we did highlight some of the problems involved in the theoretical treatment of open-shell systems within DFT. The presence of an unpaired electron induces the self-interaction error, which makes the accurate determination of dissociation energies very difficult. The incorrect description of the delocalised exchange hole produces fictitious transition states on potential energy curves and, for some systems, the dissociated atoms may end up lower in energy than the bound dimer. Furthermore, the onset of multireference character with increasing distance between the dissociating atoms results in severe SCF convergence problems. It appears that no current theoretical framework is completely free of these errors, although

Becke's real-space approach to nondynamical correlation [133, 134] looks promising.

Finally, we found that the charge-transfer interaction in the organophosphonate family increased on going from chlorine to iodine. Our theoretical predictions were supported by solid-state UV-vis data and are consistent with the findings of other authors for similar systems. Overall, it would be interesting to repeat the theoretical program of this chapter for a different choice of metal ion, *i.e.* by substituting the copper ion with another metal ion, keeping all other geometric parameters constant.

Chapter 6

Beyond DFT: The Quantum Monte Carlo Method

6.1 Introduction

The electron-electron interaction is the principal source of approximation in most commonly used theories of electronic structure, such as Hartree-Fock theory and DFT. In contrast, the form of the electron-electron interaction is not approximated at all in Quantum Monte Carlo (QMC) techniques and hence a direct treatment of quantum many-body effects is possible. Many-body effects are often required for the accurate description of quantum phenomena: van der Waals interactions and systems in which the electrons are strongly correlated,[1] to cite just two examples.

Weak physical interactions, such as van der Waals (dispersion) and hydrogen bonding, are difficult to describe using conventional quantum chemistry methods. Accurate descriptions require the use of very large basis sets in combination with high-level techniques, such as Configuration Interaction and Coupled Cluster, which are slowly convergent and scale badly with system size (N^8). QMC methods have several important advantages over conventional wavefunction techniques:

[1] For example, high temperature superconductors and colossal-magneto-resistance materials.

- The cost of a QMC calculation scales like the cube of the system size[2] so calculations on large systems are feasible;
- QMC methods can treat both finite and periodic systems, unlike Coupled Cluster, which is currently only applicable to small molecules;
- The QMC algorithm is trivially parallelisable and this makes QMC simulations ever more feasible as the availability of powerful supercomputers increases;
- Much more can be obtained from QMC calculations than just highly accurate 'numbers'. For example, QMC methods have been used to calculate exchange-correlation (XC) holes in both the homogeneous electron gas [143] and solids. [144] Analysis of the properties of the XC hole not only provides information about the suitability of current functionals, but also aids in the design of new, more accurate functionals.

It is also worth noting that QMC was the method used by Ceperley and Alder [42] to calculate the correlation properties of the homogeneous electron gas to a very high accuracy. These simulations subsequently lead to the development of the LDA.

In this chapter, we go beyond the Density Functional Theory approximation by using Quantum Monte Carlo to investigate the hydrogen bonding interaction in water dimer. In the following sections we discuss the two QMC methods used in this work, Variational Monte Carlo (VMC) and Diffusion Monte Carlo (DMC), while Section 6.3 describes how we used VMC and DMC to calculate the dissociation energy of water dimer.

6.2 Quantum Monte Carlo Methods

There are several methods that can be described as Quantum Monte Carlo. These are all based on an ability to randomly sample the wavefunction in such a way that characteristic system quantities, *e.g.* energies, can be evaluated. Let us briefly review the basic ideas behind the Monte Carlo technique.

Systems with a large number of degrees of freedom are common in physics and chem-

[2]Some parts of the algorithm can be made to scale quadratically.

istry; the many-electron problem discussed in Chapter 2 is just one example. In order to find the properties of such systems, we are normally required to evaluate integrals of very high dimension. Conventional quadrature techniques (Simpson's rule, Gaussian quadrature) can accurately evaluate one-dimensional integrals but fail when the number of dimensions becomes large. This is because the error in the quadrature falls off increasingly slowly with increasing dimension. Conversely, the uncertainty in Monte Carlo quadrature is independent of the number of dimensions, hence Monte Carlo methods provide an efficient way to evaluate many-dimensional integrals.

Monte Carlo quadrature involves two basic operations: the generation of points randomly distributed over the integration volume with a specified distribution $w(\mathbf{x})$; then evaluation of the function f/w at these points, where f is the integral to be evaluated. The fundamental advantage of Monte Carlo methods is that $w(\mathbf{x})$ is chosen such that the generation of points can be biased so they are concentrated in areas where the value of the integral is large. Thus Monte Carlo techniques typically require the evaluation of the integral at far fewer points than conventional quadrature. This can usually only be achieved with conventional quadrature techniques by finding an appropriate change of variable, which requires knowing what the integral looks like.

In quantum mechanics, the distribution we are interested in is $|\Psi(\mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r}_N)|^2$, typically a complicated multi-dimensional function. The Metropolis rejection algorithm [145] can be used to generate a sequence of points (collectively called a random walk) to sample this distribution. In QMC techniques, each point is a $3N$ dimensional vector of electron co-ordinates, $\mathbf{R} = \{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N\}$. A particular value of $\mathbf{R}$ is sometimes called a walker or configuration in the QMC literature. The main advantages of the Metropolis method are that it can be used to sample arbitrarily complex distributions and that it requires no knowledge of the normalisation, which is usually unknown or at least very complicated. Let us consider this algorithm in more detail.

Suppose we wish to use the Metropolis algorithm to sample some known distribution, $\rho(\mathbf{R})$. It can be shown that the sampling is most easily accomplished if the sequence of points form a Markov chain. Markov chains have the important property that the state-to-state transition probabilities depend only on the current state of the system,

not on how it got there. As a result, a Markov process may be completely specified by choosing appropriate values of the transition probabilities, $P(\mathbf{R} \rightarrow \mathbf{R}')$. The Metropolis method works by choosing transition probabilities in such a way that the points on the random walk sample the required probability distribution.

The generation of sampling points begins by starting the walker at some random position, $\mathbf{R}$. A trial move is made to a nearby point $\mathbf{R}'$, chosen from some probability density function, $T(\mathbf{R} \rightarrow \mathbf{R}')$. The move is accepted with probability

$$A(\mathbf{R} \rightarrow \mathbf{R}') = \text{Min}\left(1, \frac{T(\mathbf{R}' \rightarrow \mathbf{R})\rho(\mathbf{R}')}{T(\mathbf{R} \rightarrow \mathbf{R}')\rho(\mathbf{R})}\right). \tag{6.1}$$

If the move is rejected, $\mathbf{R}$ becomes the next point on the walk, *i.e.* rejected points are not discarded. Only the ratio of probabilities is calculated, so there is no need to know the normalisation of $\rho(\mathbf{R})$. Noting that $P(\mathbf{R} \rightarrow \mathbf{R}') = T(\mathbf{R} \rightarrow \mathbf{R}')A(\mathbf{R} \rightarrow \mathbf{R}')$, it is straightforward to show that

$$\frac{P(\mathbf{R} \rightarrow \mathbf{R}')}{P(\mathbf{R}' \rightarrow \mathbf{R})} = \frac{\rho(\mathrm{R}')}{\rho(\mathrm{R})}. \tag{6.2}$$

This condition is known as detailed balance. In fact, our earlier choice of $A(\mathbf{R} \rightarrow \mathbf{R}')$ was not unique. Other choices are allowed providing that they too result in detailed balance.

In a real simulation, there will be many walkers all simultaneously executing random walks according to the above rules. Each walker moves in accordance with the rules outlined above to generate its own chain. After an equilibration period, the distribution of walkers will settle down to an equilibrium distribution, $N(\mathbf{R})$. Let us show that this distribution is proportional to $\rho(\mathbf{R})$, the distribution we are trying to sample. In equilibrium, the rate at which particles arrive at $\mathbf{R}$ must exactly balance the rate at which they leave, implying that $N(\mathbf{R})$ satisfies

$$\sum_{\mathbf{R}'} P(\mathbf{R}' \rightarrow \mathbf{R})N(\mathbf{R}') = \sum_{\mathbf{R}'} P(\mathbf{R} \rightarrow \mathbf{R}')N(\mathbf{R}). \tag{6.3}$$

Combining this equation with the detailed balance condition, (6.2), we find that

$$\sum_{\mathbf{R}'} P(\mathbf{R} \rightarrow \mathbf{R}') \left(N(\mathbf{R}) - \frac{\rho(\mathbf{R})}{\rho(\mathbf{R}')} N(\mathbf{R}') \right) = 0, \tag{6.4}$$

from which it immediately follows that

$$\frac{\rho(\mathbf{R})}{\rho(\mathbf{R}')} = \frac{N(\mathbf{R})}{N(\mathbf{R}')}, \tag{6.5}$$

as required.

We have shown that the Monte Carlo technique for the evaluation of many-dimensional integrals, in combination with the Metropolis algorithm, is a powerful and efficient strategy which can be applied to a broad class of problems in both statistical and quantum mechanics.

A brief treatment of the two QMC methods used in this work follows. For a more comprehensive treatment, the reader is referred to the review by Foulkes, *et al.* [146]

6.2.1 Variational Monte Carlo

The simplest QMC method for the evaluation of ground state energies is Variational Monte Carlo (VMC). This technique relies on a combination of the variational principle (which provides a rigourous upper bound to the ground state energy) and Monte Carlo quadrature (to evaluate the appropriate integral). In VMC, the integral we are interested in is

$$E = \frac{\int \psi^*(\mathbf{R}) \hat{H}(\mathbf{R}) \psi(\mathbf{R}) \, d\mathbf{R}}{\int \psi^*(\mathbf{R}) \psi(\mathbf{R}) \, d\mathbf{R}}, \tag{6.6}$$

where $\mathbf{R}$ is as usual a $3N$ dimensional vector of electron co-ordinates, *i.e.* the walker. The main ideas behind the VMC method can be illustrated with a simple example. Let us consider the integral of some abitrary function $f(x)$,

$$F = \int_a^b f(x) dx. \tag{6.7}$$

If the integrand is sampled at N points, then the value of the integral can be approximated as

$$F_N = \frac{b-a}{N}\sum_{i=1}^{N} f(X_i). \tag{6.8}$$

The sample points, $\{X_i\}$, are drawn from some appropriate probability distribution. The RHS of (6.8) is called the estimator of the integral in equation (6.7) and its expected value is equal to the value of the integral.

It was mentioned earlier that a fundamental advantage of Monte Carlo methods is that the generation of sampling points can be biased so they are concentrated in areas where the value of the integral is large. This idea is known as importance sampling. To achieve this, we introduce a new function $w(x)$, which is a good approximation to $f(x)$. Now suppose that we can generate random numbers from the normalised function,

$$p(x) = \frac{w(x)}{\int_a^b w(x)dx}. \tag{6.9}$$

Since $w(x)$ is chosen to approximate $f(x)$, sample points drawn from $p(x)$ have the greatest concentration in areas where $f(x)$ is large. Finally, a new function, $g = f(x)/p(x)$ is defined, and (6.7) is re-written as

$$F = \int_a^b g(x)p(x)dx \approx \frac{1}{N}\sum_{i=1}^{N} g(X_i). \tag{6.10}$$

In VMC, a parameterised trial wavefunction $\psi_T(\mathbf{R})$ is used as an approximation to the true many-body wavefunction. Using the variational principle, an upper bound to the ground state energy is found to be

$$E_0 \leq \frac{\int \psi_T^*(\mathbf{R})\hat{H}(\mathbf{R})\psi_T(\mathbf{R})\, d\mathbf{R}}{\int \psi_T^*(\mathbf{R})\psi_T(\mathbf{R})\, d\mathbf{R}}. \tag{6.11}$$

It is then helpful to re-write equation (6.11) in the same form as equation (6.10):

$$E_{\mathrm{VMC}} = \int E_L(\mathbf{R}) \left(\frac{\psi_T^*(\mathbf{R})\psi_T(\mathbf{R})}{\int \psi_T^*(\mathbf{R})\psi_T(\mathbf{R})\, d\mathbf{R}} \right) \tag{6.12}$$

where the first term on the RHS is the local energy and is defined as

$$E_L(\mathbf{R}) = \frac{\hat{H}\psi_T(\mathbf{R})}{\psi_T(\mathbf{R})}. \tag{6.13}$$

Hence using the VMC method, the Metropolis algorithm can be used to estimate the ground state energy. It is the local energy, not the potential energy, which is computed at each of the Monte Carlo sample points. We will discuss why this is advantageous in the next section on Diffusion Monte Carlo.

Finally, the VMC energy is calculated as the local energy summed over the number of sampling points:

$$E_{\mathrm{VMC}} \approx \frac{1}{N}\sum_{n=1}^{N} E_L(\mathbf{R}_n). \tag{6.14}$$

Note the relationship between equations (6.10) and (6.14).

The accuracy of the VMC approach is severely limited by the form of the trial wavefunction: the closer ψ_T is to the true ground state wavefunction, the more accurate the estimate is. The subject of trial wavefunctions is discussed in more detail in Section 6.2.3. A more accurate, but computationally intensive, alternative to VMC is the Diffusion Monte Carlo method, described below.

6.2.2 Diffusion Monte Carlo

The starting point for the Diffusion Monte Carlo (DMC) method is the many-body Schrödinger equation written in imaginary time

$$-\frac{\partial\Psi(\mathbf{R},\tau)}{\partial\tau} = [-D\nabla^2 + V(\mathbf{R}) - E_T -]\Psi(\mathbf{R},\tau), \tag{6.15}$$

where $\tau = it$, $D = 1/2$ and E_T is an adjustable energy shift. Equation (6.15) is a standard diffusion equation in the real variable τ. The many-body wavefunction, $\Psi(\mathbf{R},\tau)$, represents the density of diffusing walkers. If we ignore the $[V(\mathbf{R}) - E_T]$ term, (6.15) reduces to a diffusion equation with diffusion constant $D = 1/2$ and describes a population of walkers diffusing in $3N$-dimensional configuration space. If only the $[V(\mathbf{R}) - E_T]$ term is present, (6.15) is a rate equation which describes a branching process such as radioactive decay. The entire equation can be simulated using random

walks, where the density of diffusing walkers increases in regions of low potential energy and decreases in regions of high potential energy. The energy shift, E_T, is occasionally adjusted to keep the total walker population roughly constant.

Using random walks with branching to solve equation (6.15) is usually very inefficient. The rate of branching is controlled by $V(\mathbf{R})$, which can fluctuate wildly and even diverge, *e.g.* when two charged particles coincide. This subsequently leads to large fluctuations in the walker population and hence increased statistical uncertainties. These numerical instabilities can be minimised by using importance sampling. Here as with VMC, we can multiply the exact ground state wavefunction, $\Psi_0(\mathbf{R},\tau)$, by a trial function, $\Psi_T(\mathbf{R})$, to form a new probability distribution:

$$f(\mathbf{R},\tau) = \Psi_0(\mathbf{R},\tau)\Psi_T(\mathbf{R}). \tag{6.16}$$

Equation (6.15) can now be re-written as

$$-\frac{\partial f(\mathbf{R},\tau)}{\partial \tau} = -D\nabla^2 f(\mathbf{R},\tau) + (E_L(\mathbf{R}) - E_T)f(\mathbf{R},\tau) + \nabla\cdot(\mathbf{v}_D(\mathbf{R})f(\mathbf{R},\tau)) \tag{6.17}$$

where $E_L(\mathbf{R})$ is the local energy as defined for the VMC method (6.12) and $\mathbf{v}_D(\mathbf{R})$ is the $3N$-dimensional drift velocity defined by

$$\mathbf{v}_D(\mathbf{R}) = \nabla \ln|\Psi_T(\mathbf{R})| = \nabla\Psi_T(\mathbf{R})/\Psi_T(\mathbf{R}). \tag{6.18}$$

The branching rate is now controlled by E_L, rather than $V(\mathbf{R})$. For a good choice of Ψ_T, the local energy is roughly constant and the walker population is accordingly stabilised. The drift velocity serves to increase the density of walkers in regions where Ψ_T is large and also helps to satisfy the fixed-node constraint, which we will discuss later. Importance sampling thus greatly improves the efficiency of the DMC method.

The distribution of (6.17) in the long-time limit is proportional to the ground state wavefunction. However, the Monte Carlo method is poorly suited for solving differential equations directly. The equation in question must first be converted to integral form and this involves determining the Green's function. The integral equation equivalent

to (6.17) is

$$f(\mathbf{R}', \tau, \Delta\tau) = \int G(\mathbf{R} \rightarrow \mathbf{R}', \delta\tau) f(\mathbf{R}, \tau) d(\mathbf{R}) \tag{6.19}$$

where $G(\mathbf{R} \rightarrow \mathbf{R}', \Delta\tau)$ is the Green's function and can be interpreted as a transition probability density for the evolution of the walkers in time $\Delta\tau$. The explicit expression for the exact Green's function for interacting particles is not known, except in a few special cases. We must therefore introduce some approximations.

The full Hamiltonian contains both kinetic ($\hat{T}$) and potential energy ($\hat{V}$) terms. The Trotter product formula for an exponential of a sum of operators can be used to separate $\hat{T}$ and $\hat{V}$ to form a new expression:

$$\begin{aligned} G(\mathbf{R} \rightarrow \mathbf{R}', \Delta\tau) &= \langle \mathbf{R} | e^{-\Delta\tau(\hat{T}+\hat{V}-E_T)} | \mathbf{R}' \rangle \\ &\approx e^{-\Delta\tau[V(\mathbf{R})-E_T]/2} \langle \mathbf{R} | e^{-\Delta\tau\hat{T}} | \mathbf{R}' \rangle \\ &\quad \times e^{-\Delta\tau[V(\mathbf{R}')-E_T]/2}. \end{aligned} \tag{6.20}$$

For short times $\Delta\tau$ we may assume that both the local energy and drift velocity are constant, independent of $\mathbf{R}$. The short-time approximation to the Green's function for small τ is given by [147]:

$$\begin{aligned} G(\mathbf{R} \rightarrow \mathbf{R}', \Delta\tau) &\approx (2\pi\Delta\tau)^{3N/2} \exp\left[-\frac{(\mathbf{R}-\mathbf{R}')^2}{2\Delta\tau}\right] \\ &\quad \times \exp[-\Delta\tau[V(\mathbf{R} + V(\mathbf{R}') - 2E_T]/2]. \end{aligned} \tag{6.21}$$

The error induced by this approximation is proportional to $(\Delta\tau)^3$ and arises because the operators, $\hat{T}$ and $\hat{V}$, do not commute. In practice, the effect of this error (the so-called time-step bias) can be minimised by performing a calculation at a very small value of $\Delta\tau$ or by performing several calculations at different values of $\Delta\tau$ and then extrapolating to zero time.

Repeated operation by $G(\mathbf{R} \rightarrow \mathbf{R}', \Delta\tau)$ on the distribution $f(\mathbf{R}, \tau)$ results in an equilibrium distribution proportional to $\Psi_T(\mathbf{R})\Psi_0(\mathbf{R})$. The walkers sample the local energy at each point of their random walk through configuration space, in the same manner as for the VMC method. The average of these local energy values is the mixed estimator

of the ground state energy,

$$E_{DMC} = \int \left(\frac{\hat{H}\Psi_T}{\Psi_T} \right) \left(\frac{\Psi_T \Psi_0}{\int \Psi_T \Psi_0 \, d\mathbf{R}} \right) d\mathbf{R}. \tag{6.22}$$

The presence of the exact ground state wavefunction in (6.22) ensures that the mixed estimator delivers the exact ground state energy.

The DMC method is more accurate than VMC because in DMC, the trial function is only used as the starting point for the simulation and branching then corrects the distribution $|\Psi_T|^2$ in regions where it is a poor approximation to Ψ_0. The DMC method is thus much less sensitive to the choice of trial wavefunction than VMC. The subject of trial wavefunctions will be discussed in more detail in the following section.

There is one more important point concerning DMC which needs to be considered. The diffusion equation formulation of DMC requires that the density of diffusing walkers, represented by the many-body wavefunction, be everywhere positive and not change sign. However, Fermi statistics dictate that many-electron wavefunctions be antisymmetric under exchange of electron co-ordinates, *i.e.* many-electron wavefunctions have both positive and negative regions. The most common method used to cope with wavefunction antisymmetry is the fixed-node approximation. [148] In this approach, the exact many-body wavefunction is constrained to have the same nodes as the trial wavefunction.[3] The nodes of the trial wavefunction then define the surface on which the exact wavefunction vanishes and equation (6.17) is solved separately in each nodally bounded volume of Ψ_T. Because we are forced to use an approximate Green's function, walkers will occasionally attempt to cross nodes and this can result in significant bias. Importance sampling is very helpful in minimising node crossings because the drift velocity grows as a walker approaches a node and carries it away.

Since the approximate nodal surface is usually obtained from a variational calculation, fixed-node DMC energies are accordingly variational, *i.e.* the DMC energy is always greater than or equal to the exact ground state energy for a particular nodal surface.

[3] The nodes of the trial wavefunction, obtained from a DFT or HF calculation, are known; the nodes of the exact many-body wavefunction are not.

If the trial nodal surface were exact, the DMC method would deliver the exact ground state energy. The fixed-node approximation is the only approximation made in the DMC method and typically leads to an error of about 5%.

6.2.3 The Trial Wavefunction

In Sections 6.2.1 and 6.2.2 we have made repeated mention of the trial wavefunction, an integral component of both the VMC and DMC methods. The final accuracy of a QMC simulation is determined by the quality of the trial wavefunction. A good trial wavefunction will also improve the efficiency of the simulation and minimise statistical uncertainties and numerical instabilities. There is much to be said concerning the subject of trial wavefunctions so we will only consider them briefly here. The interested reader is referred to Foulkes, *et al* [146] for more information.

The type of trial wavefunction most commonly used in QMC simulations is of the Slater-Jastrow form:

$$\Psi(\mathbf{X}) = \exp\left[J(\mathbf{X})\right] D(\mathbf{X}) \tag{6.23}$$

where $J(\mathbf{X})$ is the Jastrow correlation function [149], $D(\mathbf{X})$ is a Slater determinant, $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$, and $\mathbf{x}_i = \{\mathbf{r}_i, \sigma_i\}$ are the space and spin co-ordinates of electron i. The Jastrow function is a totally symmetric nodeless function which explicitly contains the inter-electronic co-ordinates and has the basic form:

$$J(\mathbf{X}) = \sum_{i=1}^{N} \chi(\mathbf{x}_i) - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1,(j\neq i)}^{N} u(\mathbf{x}_i, \mathbf{x}_j). \tag{6.24}$$

The u function is a two-body term which describes correlations between electrons while the χ function describes correlations between electrons and nuclei. The Jastrow function used in this work [150] also contains three-body electron-electron-nucleus terms centred on the nuclei. It is computationally advantageous to remove the spin variables from the Slater-Jastrow wavefunction and replace the single Slater determinant by a product of up-spin and down-spin orbitals,

$$\Psi(\mathbf{R}) = \exp\left[J(\mathbf{R})\right] D_{\uparrow}(\mathbf{r_1}, \ldots, \mathbf{r_{N\uparrow}}) D_{\downarrow}(\mathbf{r_{N\uparrow+1}}, \ldots, \mathbf{r_N}) \tag{6.25}$$

where $\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ denotes the spatial co-ordinates of all the electrons. The orbitals are usually obtained from a HF or DFT calculation.

The accuracy of the VMC and DMC methods depends crucially on the properties of the trial wavefunction. The repeated evaluation of the wavefunction is also the most costly part of a QMC simulation. The chosen trial wavefunction must thus be accurate, compact and easy to evaluate. One way to achieve this is to optimise the trial wavefunction by minimising its variance.[4] The usual procedure involves using VMC to generate a set of electron configurations $\{\mathbf{R}\}$ distributing according to $[\Psi^{\{\alpha_0\}}(\mathbf{R})]^2$, where $\{\alpha_0\}$ is some fixed parameter set (the Jastrow function contains the optimisable parameters). The energy variance is then minimised over this parameter set. Since the variance landscape depends on the distribution of configurations, several configuration generation/variance minimisation iterations are performed and the wavefunction that produces the lowest variational energy is chosen. The optimisation procedure used in this work is that of Drummond and Needs. [151]

Assumptions and Limitations

The fixed-node approximation is the major approximation made in the DMC method, that is, the wavefunction can be solved exactly within the nodal pockets but the nodes themselves are fixed by the trial wavefunction. The main advantage of this method is that high-accuracy results can be routinely obtained (given a good trial wavefunction). In addition, there is no basis set superposition error and the results are only weakly dependent on the form of the wavefunction used to represent the nodal surface.

A significant technical disadvantage is the present inability of QMC to calculate forces and therefore perform geometry optimisation and dynamics. This is a significant factor in preventing QMC from becoming a routine technique. Although some progress has been made, the calculation of forces in QMC remains unrealistic for all but the smallest systems. On a more practical side, the lack of generally available programs to perform QMC calculations means that the method is not as accessible as other standard electronic structure techniques and its widespread use and testing of broader applicability

[4] Energy minimisation schemes exist but it is easier to find stable algorithms for variance minimisation. Furthermore, unlike the energy, the variance has a known lower bound of zero.

is still some time ahead. However, accurate geometries may be obtained by methods which do not give particularly accurate energies (DFT) and then one can use QMC to obtain a high-accuracy, single-point energy.

Having noted the above, we should remember that Coupled Cluster methods are even more computationally demanding than QMC and one is often practically restricted to very small systems or forced to forgo a geometry optimisation altogether and use a less accurate method to obtain the optimum geometry, as described above. For larger systems, even single-point calculations may not be possible. In this respect, the Coupled Cluster method offers no advantage over QMC.

6.3 Quantum Monte Carlo Calculations of the Dissociation Energy of Water Dimer

The description of the structure and energetics of assemblies of water molecules is one of the most fundamental problems in electronic structure theory, as this knowledge is vital to the understanding of water in all its physical states. Indeed, water dimer has probably been one of the most intensively studied hydrogen bonded systems, both theoretically and experimentally. Despite its apparent simplicity, accurate theoretical descriptions of water dimer have typically required the use of very large basis sets in combination with sophisticated wavefunction techniques, such as Configuration Interaction (CI) and Coupled Cluster (CC). As we discussed earlier, while these methods are highly accurate, they are very computationally intensive. They also suffer from basis set superposition and basis set incompleteness errors. Density Functional Theory (DFT) represents a computationally economical choice however, results for water dimer vary significantly depending on the functional used to perform the calculations. DFT is also plagued by the same basis set errors which occur in CI and CC calculations. Water dimer would thus seem an ideal test case for QMC.

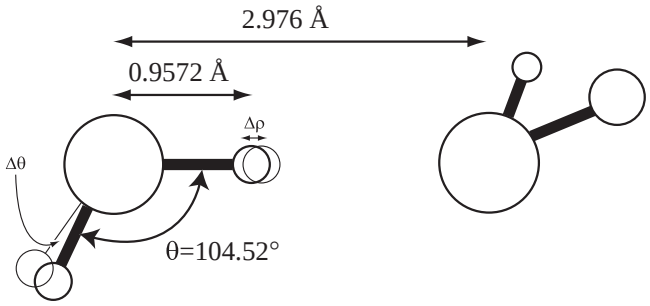

Figure 6.1: The linear configuration of the water dimer. The equilibrium bond angle of the subunit water monomer of α=104.52° is shown, as is the R_{OH} bond length of 0.9572 Å. The hydrogen bond length of $R_{O\cdots O}$=2.976 Å is indicated, along with structural deformations that may occur upon dimerisation: $\Delta\theta$ denotes angular relaxation and $\Delta\rho$ denotes OH bond extension. Structural deformations were ignored in this work due to their negligible influence on the total energy. Image kindly supplied by Richard Needs.

6.4 Computational Details

6.4.1 System Selection and Trial Wavefunction Construction

We have performed calculations on water monomer at the experimentally determined equilibrium geometry, [95] given by r_{OH}=r'_{OH}=0.9572Å and $\measuredangle_{HOH}$=104.52°. This geometry has been used in previous calculations on water monomer. [152]

Experiments have determined that the equilibrium conformation of water dimer is a linear configuration with C_s symmetry, [106, 153, 107] as shown in Figure 6.1. The oxygen-oxygen (O–O) distance was measured to be 2.976 Å. This was corrected to 2.946 Å after anharmonic vibrational effects were taken into account. Subsequent theoretical calculations estimated R_{O-O} to be 2.943 Å [154] and 2.953 Å [155]; recent estimates suggest an even shorter R_{O-O} of 2.912 Å [105]. However, Åstrand and coworkers [156] have suggested (as pointed out by Klopper, *et al* [105]), that in theoretical determinations of the geometry, the out-of-plane motion of the hydrogen atom involved in the hydrogen bond should be considered, in addition to the in-plane stretching frequency. Their calculations, with the effects of out-of-plane motion included, yielded an R_{O-O} of 2.982 Å. All of our calculations have therefore been performed on the experimentally determined structure with a R_{O-O} of 2.976 Å.

Upon forming a dimer, some structural deformation occurs in the monomer subunits: there is a slight change in the angle between the inter-oxygen vector and the plane of

Table 6.1: Structure of the linear water dimer in atomic units.

Atom	x	y	z
O	-2.80630	0.00000	0.00000
H	-0.99742	0.00000	0.00000
H	-3.25982	0.00000	-1.75111
O	2.80630	0.00000	0.00000
H	3.35989	1.43046	0.95885
H	3.35989	-1.43046	0.95885

the proton-accepting monomer and the bond angle for the proton acceptor widens by about 0.5°. However, these changes are small and thus it is reasonable to assume that the deformation energy is negligible compared to the intermolecular binding energy. We have thus chosen to ignore any angular relaxations. The co-ordinates we have used for the linear water dimer are given in Table 6.1.

We have used trial wavefunctions of the Slater-Jastrow form, as discussed in Section 6.2.3. The Slater determinants, $D(\mathbf{X})$, have been obtained from Hartree-Fock approximation (HF) and Density Functional Theory (B3LYP functional) calculations using the Gaussian03 code. [87] Our QMC calculations have thus been performed using two different approximate nodal surfaces. The error induced by the fixed-node approximation should largely cancel when energy differences, *e.g.* dissociation energies, are calculated. The use of two different approximate nodal surfaces in this work allows us to assess the extent to which the DMC water dimer dissociation energy depends on the nodal surface. Our Jastrow functions contained a total of 108 adjustable parameters, which were optimised using the variance minimisation procedure described in Section 6.2.3. We used the ANO basis of Widmark, Malmqvist and Roos [157], which tests showed to give good results for water monomer and dimer. Although the trial wavefunctions contain orbitals expanded in a basis set, the energies obtained in the DMC method are less dependant on the quality of the single-particle basis than HF, VMC or CCSD(T) energies. All of our DMC calculations have been performed with a target population of 1280 configurations. Time step errors have been carefully checked. We performed DMC calculations at several time steps and extrapolated the result to zero time. All of our VMC and DMC simulations have been performed using the CASINO code. [158]

6.5 Results and Discussion

6.5.1 Water Monomer

Table 6.2 shows our results for the total energy of the water monomer and those from other QMC calculations and from CC calculations. Our HF energy is only 0.005 Ha higher than the estimated basis-set limit result [159], which demonstrates the high quality of the basis set used. [157] In Section 2, we noted that the accuracy of a VMC simulation was entirely limited by the quality of the trial wavefunction; this is reflected in our VMC-HF and VMC-B3LYP energies. While they are lower than their respective HF and B3LYP equivalents, they are still significantly higher than our DMC energies and the "exact" result. Our DMC energies with HF nodes (DMC-HF) and B3LYP nodes (DMC-B3LYP) are 1 and 2 mHartree (mHa), respectively, below Manten and Lüchow's DMC result of -76.4207(2).[160] Our DMC-B3LYP energy is only 15 mHa above the "exact" result of -76.438 Ha. Our DMC-B3LYP energy is the best single-determinant DMC energy reported so far, and is only 6 mHa above the DMC result of Lüchow and Fink [161], who used a trial wavefunction containing 300 determinants. Müller and co-workers obtained an energy of -76.4373 Ha using the CCSD(T)-R12 method which, although not variational, is the best calculated result reported to date.

6.5.2 Water Dimer

The accuracy of the experimental estimate for D_e has recently been brought into question, perhaps in the wake of a series of high-level calculations, which consistently favour the lower limit. The enthalpy of association of water dimer was measured in thermal conductivity experiments [106] to be -3.59±0.5 kcal/mol. The zero-point energy (ZPE) contribution was calculated theoretically to arrive at a final estimate of -5.44±0.7 kcal/mol. Mas and co-workers have argued [164] that the ZPE was calculated using a very small (by today's standards – 4-31G) basis set and that the frequencies used differ from the current best estimates by more than 50% in some cases. The Symmetry Adapted Perturbation Theory (SAPT) calculations of Mas, *et al* predict a ZPE of 1.81 kcal/mol , which results in an estimated dissociation energy of -5.00±0.7 kcal/mol.

Table 6.2: DMC total energies of H_2O and comparisons with various *ab initio* methods. The entry 76.327(5), for example, denotes a standard error in the mean of 5 units in the last figure, *i.e.* 76.327±0.005.

Method	Total Energy (Hartree)
HF[1]	-76.063
HF limit[2]	-76.068
B3LYP[1]	-76.469
VMC-HF[1]	-76.327(5)
VMC-B3LYP[1]	-76.334(5)
CCSD(T)/aug-cc-pV5Z[2]	-76.370293
DMC[3]	-76.4207(2)
DMC-HF[1]	-76.4219(1)
DMC-PNO-CI[4]	-76.429(1)
DMC-B3LYP[1]	-76.4230(1)
CCSD(T)-R12[5]	-76.4373
"Exact"[6]	-76.438

[1] This work.
[2] Ref. [78].
[3] Ref. [160].
[4] With Pair Natural Orbital–Configuration Interaction trial functions, Ref. [161].
[5] Ref. [162].
[6] Ref. [163].

Table 6.3: DMC dissociation energy of water dimer and comparisons with various *ab initio* methods.

Method	D_e (kcal/mol)
SAPT-5s[1]	-4.86
CCSD(T)-Extrapolated[2]	-5.02±0.05
DMC[3]	-5.66±0.20
DMC-HF[4]	-5.02±0.18
DMC-B3LYP[4]	-5.21±0.18
Experiment,'corrected'[5]	-5.00±0.7
Experiment[6]	-5.44±0.7

[1] Ref. [164].
[2] Ref. [105].
[3] Ref. [165].
[4] This work.
[5] Ref. [164].
[6] Ref. [106].

Table 6.3 shows our results for the dissociation energy of water dimer. Our DMC energy with HF nodes (DMC-HF) compares very well with the CCSD(T) result of Klopper *et al* [105] and the corrected experimental value. We obtain a lower binding energy for the wavefunction with B3LYP nodes (DMC-B3LYP – -5.21±0.18 kcal/mol), although this value is still within the error bars of the DMC-HF result. It should be noted that Klopper and co-workers used a theoretically optimised geometry while we have used the experimentally determined equilibrium geometry. We also note that the DMC energy of Diedrich *et al* [165] was obtained from a pseudopotential DMC calculation. Figure 6.2 shows the convergence of the dissociation energy with DMC time-step for the wavefunction with HF nodes. The time-step error essentially completely cancels at small values of the time-step.

We have attempted to estimate the effect on the dissociation energy due to a change in the O$\cdots$O bond length. Klopper, *et al* calculated dissociation energies for a CC optimised structure (R_{O-O}=2.912 Å) and a MP2 optimised one (R_{O-O}=2.910 Å). [105] The difference in dissociation energy between the two structures was approximately 0.02 kcal/mol. They also estimated an error bar of ±0.05 Å on the computed R_{O-O} for the CC optimised structure, the upper limit of which is within ∼0.01 Å of the bond length we have used. At larger values of R_{O-O}, the difference in dissociation energy due to a change in the bond length will be less pronounced. Hence, we conclude that the

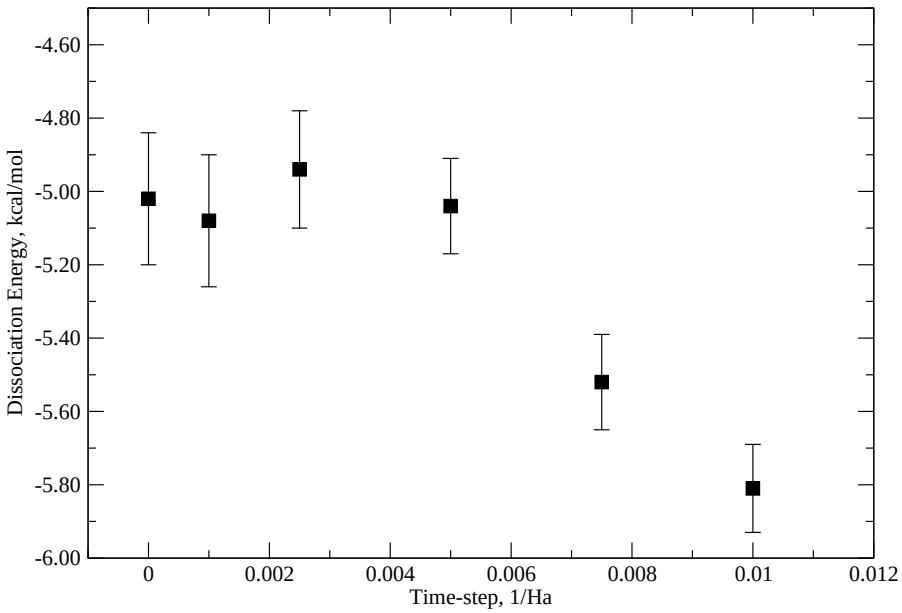

Figure 6.2: Time-step dependance of the dissociation energy of water dimer for a DMC wavefunction with HF nodes.

effect on the dissociation energy of using a different geometry to that of Klopper and co-workers is contained within our DMC error bars.

The DMC and CC methods are inherently very different to one another, so the good agreement between the two methods concerning dissociation energies is encouraging. It is also interesting to note that all three high-level results quoted here (CC, DMC-HF and DMC-B3LYP) favour the lower limit of the experimental estimate and are in good agreement with the corrected experimental estimate of -5.00±0.7 kcal/mol. This lends further support to the validity of the corrected value.

In Section 6.2.2 we noted that the fixed-node approximation is the major approximation made in the DMC method and that if the exact nodal surface was known, DMC would return the exact ground state energy. This task is complicated by the fact that currently one can only determine the quality of a particular nodal surface by performing an expensive DMC calculation, which in any case does not give information about how to improve the nodal surface. Several authors have investigated the effect of the single-particle orbitals on the nodal error, with mixed results. Grossman [166] calculated the atomisation energy of P_2 using a single Slater determinant built from five different

kinds of orbitals: Hartree-Fock, Natural Orbitals (NO), LDA, BPW91 and B3LYP. He found that the DFT orbitals returned the best result, followed by NO and HF. These results are consistent with the findings of Wagner and Mitas [167], who calculated the binding energy of MnO and TiO using HF and B3LYP orbtals; they also found that the B3LYP orbitals produced more accurate DMC binding energies than HF orbitals. In contrast, Bressanini and co-workers [168] found that, for the He_2^+ system, Complete Active Space (CAS) and HF orbitals produced the most accurate DMC results. DFT orbitals, generated from several different functionals, were found to perform very poorly although the authors do not state exactly which functionals were tested. It appears that the behaviour of the nodal error is very much dependant on the system being studied. Our results suggest that the water dimer dissociation energy is relatively insensitive to the choice of nodal surface although we obtained a lower water monomer and dimer total energy with B3LYP nodes.

6.6 Conclusions

In this chapter, we performed all-electron HF, DFT, and QMC calculations of the total energy of water monomer and dimer and have compared our results for the dissociation energy of water dimer with both CCSD(T) calculations and the available experimental data. We predict dissociation energies for the dimer of -5.02±0.18 kcal/mol (DMC-HF) and -5.21±0.18 kcal/mol (DMC-B3LYP), in good agreement with the best reported Coupled Cluster dissociation energy of -5.02±0.05 kcal/mol and the alternative experimental dissociation energy of -5.00±0.7 kcal/mol. Our results also indicate that the water dimer dissociation energy is relatively insensitive to the nodal surface of the trial wavefunction, although we noted that the behaviour of the *nodal error* is sensitive to the system being studied.

The R_{O-O} distance in water dimer remains a point of contention among theorists. The calculation of forces within QMC is very challenging, though progress has been made in this area. In the case of water dimer, only the $O-O$ distance needs to be optimised and this can be achieved through a series of total energy calculations. It would thus be interesting to calculate the DMC energy for $O-O$ distances in the range 2.90 Å to

3.00 Å to see where the 'DMC minimum' is. Unfortunately, such a task was too time-consuming to be attempted here.

While it was possible to perform all-electron DMC calculations for water dimer, it was very computationally demanding (with regards to time only; DMC calculations are not generally memory-intensive). One can envisage, that for a larger molecular system, the computational time demands would become increasingly prohibitive. Much work has been conducted concerning the use of pseudopotentials in QMC [169, 170], in both finite and periodic systems. In fact, pseudopotential QMC calculations are already the norm for condensed matter applications and have been for some time. Pseudopotentials allow larger time steps to be used than are possible for all-electron QMC calculations, thereby reducing the computational effort required without compromising the attainable accuracy. Hence, QMC with pseudopotentials should be a very powerful technique for the accurate calculation of binding energies in weakly-bound systems, particularly those too large to be treated with Coupled Cluster techniques.

Chapter 7

Conclusions

In the current work, we have used a number of theoretical methods to investigate the structures and energetics of the molecular components which comprise the organophosphonate ionic molecular crystals. We have also tested the applicability of Quantum Monte Carlo methods to weakly-bound molecular systems, taking water dimer as our test case. The uniting theme of this thesis has been the investigation of intra- and intermolecular interactions, which span the energy range from weak to strong, in different environments using a variety of theoretical methods. In this final chapter, we summarise our work and consider what we have learnt about interactions in the organophosphonate family and interactions in general.

Our study of hydrogen bonding interactions in the phenylphosphonic acid dimers required us to first design both an accurate and efficient methodology. As part of this process, in Chapter 3, we selected a set of well-characterised hydrogen-bonded systems and computed the structures and binding energies with a number of exchange-correlation functionals in combination with two different types of basis set – Gaussian-type and numerical. The numerical basis sets were found to provide the required efficiency while the PBE functional produced accurate geometries.

The findings of the study in Chapter 3 were applied to our investigation of the phenylphosphonic acid dimers in Chapter 4. The optimised structures of each of our

model systems were in good agreement with the corresponding experimental structures, confirming the suitability of our chosen methodology. The calculated binding energies for both the substituted and unsubstituted dimers were consistent with those found for other ionic hydrogen bonded systems. The hydrogen bonding network formed by the phenylphosphonic acid dimers appears to be reasonably robust to changes in the acid. We also investigated the effect of different parts of the crystal field on the geometry of a single, unsubstituted phenylphosphonic acid dimer anion. Different parts of the crystal field were found to affect the geometry of the anion in different ways, with the C-P-O-H torsion angle being the most sensitive parameter to crystal field effects. We concluded that the geometry of the anion was likely a complex product of competing intra- and intermolecular forces and crystal field effects. In the final section of Chapter 4, we considered the co-operative nature of the interactions in a chain of cation-anion pairs for each member of the organophosphonate family. The interaction energy was found to be mostly one-body (local) in nature and varied with the type of halogen attached to the copper ion.

In Chapter 5, we turned our attention to the other component of the organophosphonate system, a paramagnetic copper-halo-phenanthroline charge-transfer complex. Continuing our investigation of crystal field effects from Chapter 4, we looked at the stereochemical preferences of the Cu(I) starting material and Cu(II) product. The difference in geometry between the two complexes is due to the electronic configuration of the Cu ion, not crystal field effects (although the Cu(I) complex is more susceptible to deformation by the crystal field than the Cu(II) product). Following on from this work, we attempted to simulate the step-by-step dissociation of the $[CuI]^+$ system. Although this investigation was not successful, we were able to expose some of the pitfalls of simulating dissociating odd-electron systems: the self-interaction error, failure of local exchange-correlation functionals to describe delocalised exchange holes and the onset of multireference character with increasing inter-fragment distance (which subsequently lead to a decreasing HOMO-LUMO gap and severe SCF convergence difficulties). Finally, we compared the charge-transfer interaction for the chloro-, bromo- and iodo-forms of the organophosphonate series. Theoretical predictions were supported by experimental solid state UV-visible spectra, which showed that the charge-transfer

interaction increased on going from Cl to I.

Chapter 6 saw us take a different approach to the simulation of interactions. Taking water dimer as our test case, we tested the ability of the QMC technique to describe weakly-bound molecular systems. QMC is different to DFT, most significantly because the form of the electron-electron interaction is not approximated at all in QMC techniques. The correlation treatment is thus exact. For water monomer, we obtained a DMC-B3LYP total energy of -76.4230(1)Ha, the lowest single-determinant energy for water monomer to date. Our binding energies of -5.02±0.18 kcal/mol (DMC-HF) and -5.21±0.18 kcal/mol (DMC-B3LYP) are in excellent agreement with the CCSD(T) result of Klopper, *et al* and the corrected experimental binding energy. We noted that all-electron QMC for molecular (finite) systems larger than water dimer may become prohibitively expensive in terms of the computer time required. However, QMC with pseudopotentials should be a very powerful technique for the accurate calculation of binding energies for weakly-bound molecular systems, especially those which are too big to be treated by Coupled Cluster methods.

Having noted the results of Chapters 3-5, what then can we say about simulating interactions in the organophosphonate system as a whole? Our findings suggest that for sufficiently complex molecular solids (such as the organophosphonates), a single theoretical approach may not be able to describe all the components which comprise the material. The study of hydrogen bonding interactions between the phenylphosphonic acid dimers required us to pay careful attention to the choice of functional and basis set. In simulating the charge-transfer interaction in the copper-halo-phenanthroline cation, we had to be mindful of a number of complicating factors: the presence of an unpaired electron, the use of a local functional to describe a delocalised system and the use of a single determinant to describe a multireference system. In the world of theoretical inorganic chemistry, one size may definitely not fit all.

Overall, the interdisciplinary approach used in this work reflects the complex and ubiquitous nature of intra- and intermolecular interactions. Our molecular crystal research touched on aspects of organic, inorganic and quantum chemistry while our QMC study of water monomer and dimer employed the tools of many-body and theoretical physics.

It is precisely this diversity which ensures that the study of interactions will be a fertile area of research for some time to come.

Bibliography

[1] G. A. Jeffrey, *An Introduction to Hydrogen Bonding* (Oxford University Press, Oxford, 1997).

[2] R. Clarke, K. Latham, C. Rix, and J. White, "Two polymorphs of bis(1,10-phenanthroline-N,N')copper(I) iodide," Acta Cryst. C **C59,** m7–m9 (2003).

[3] R. Clarke, K. Latham, C. Rix, M. Hobday, and J. White, "Novel copper materials based on the self-assembly of organophosphonic acids and bidentate amines.," CrystEngComm **7,** 28 (2005).

[4] R. Clarke, Phd thesis, RMIT University, 2002.

[5] C. K. Chiang, C. R. Fincher, Y. W. Park, A. J. Heeger, H. Shirakawa, E. J. Louis, S. C. Gau, and A. G. MacDiarmid, "Electrical conductivity in doped polyacetylene," Phys. Rev. Lett. **39,** 1098 (1977).

[6] H. Shirakawa, E. J. Louis, A. G. MacDiarmid, C. K. Chiang, and A. J. Heeger, "Synthesis of electrically conducting organic polymers: halogen derivatives of polyactylene, $(CH)_x$," J. Chem. Soc. Chem. Commun. **16,** 578 (1977).

[7] K. Hummer, P. Puschnig, and C. Ambrosch-Draxl, "Ab initio study of anthracene under high pressure," Phys. Rev. B **67,** 184105 (2003).

[8] K. Hummer, P. Puschnig, and C. Ambrosch-Draxl, "Lowest optical excitations in molecular crystals: bound excitons versus free electron-hole pairs in anthracene," Phys. Rev. Lett. **92,** 147402 (2004).

[9] K. Hummer and C. Ambrosch-Draxl, "Oligoacene exciton binding energies: their dependence on molecular size," Phys. Rev. B **71,** 081202 (2005).

[10] R. B. Aust, W. H. Bentley, and H. G. Drickamer, "Behaviour of fused-ring aromatic hydrocarbons at very high pressure," J. Chem. Phys. **41,** 1856 (1964).

[11] M. Oehzelt, R. Resel, and A. Nakayama, "High-pressure structural properties of anthracene up to 10 GPa," Phys. Rev. B **66,** 174104 (2002).

[12] I. Dance, "Inorganic intermolecular motifs, and their energies," CrystEngComm **5,** 208 (2003).

[13] G. R. Desiraju, "The supramolecular concept as a bridge between organic, inorganic and organometallic crystal chemistry," J. Mol. Struct. **374,** 191 (1996).

[14] G. R. Desiraju, "Designer crystals: intermolecular interactions, network structures and supramolecular synthons," Chem. Comm. p. 1475 (1997).

[15] G. R. Desiraju, "Crystal engineering. From molecules to materials," J. Mol. Struct. **656,** 5 (2003).

[16] A. Clearfield, "Layered phosphates, phosphites and phosphonates of group 4 and 14 metals," Comm. Inorg. Chem. **10,** 89 (1990).

[17] G. Cao, H. Hong, and T. Mallouk, "Layered metal phosphates and phosphonates: from crystals to monolayers," Acc. Chem. Res. **25,** 420 (1992).

[18] K. A. Kraus and H. O. Phillips, "Adsorption on inorganic materials. I. Cation exchange properties of zirconium phosphate," J. Am. Chem. Soc. **78,** 694 (1956).

[19] H. Jankovics, M. Daskalakis, C. P. Raptopolou, A. Terzis, V. Tangoulis, J. Giapintzakis, T. Kiss, and A. Salifoglou, "Synthesis and structural and spectroscopic characterisation of a complex between Co(II) and imino-bis(methylphosphonic acid): Gaining insight into biologically relevant metal-ion phosphonate interactions or looking at a new Co(II)-organophosphonate material?," Inorg. Chem. **41,** 3366 (2002).

[20] K. Aoki, L. C. Brosseau, and T. E. Mallouk, "Metal phosphonate-based quartz crystal microbalance sensors for amines and ammonia," Sensors and Actuators B: Chemical **14,** 703 (1993).

[21] N. A. Benedek, M. J. S. Spencer, K. Latham, and I. Yarovsky, "Hydrogen bonding in mixed ligand copper organophosphonates," Chem. Phys. Lett. **378,** 400 (2003).

[22] W. M. Latimer and W. H. Rodebush, "Polarity and ionization from the standpoint of the Lewis Theory of Valence," J. Am. Chem. Soc. **42,** 1419 (1920).

[23] A. C. Legon and D. J. Millen, "Gas-phase spectroscopy and the properties of hydrogen-bonded dimers: HCN$\cdots$HF as the spectroscopic prototype," Chem. Rev. **86,** 635 (1986).

[24] A. C. Legon and D. J. Millen, "Directional character, strength, and nature of the hydrogen bond in gas-phase dimers," Acc. Chem. Res. **20,** 39 (1987).

[25] L. Brammer, "Direct and Indirect Roles of Metal Centres in Hydrogen Bonding," in *Implications of Molecular and Materials Structure for New Technologies, NATO ASI Series E: Applied Sciences*, H. J. Allen, F.H. and G. Shields, eds., (Kluwer Academic Publishers, Dordecht, 1999), Vol. 360, p. 197.

[26] L. Brammer, E. A. Bruton, and P. Sherwood, "Understanding the behaviour of halogens as hydrogen bond acceptors," Crystal Growth & Design **1,** 277 (2001).

[27] L. Brammer, J. C. Mareque Rivas, R. Atencio, S. Fang, and C. F. Pigge, "Combining hydrogen bonds with coordination chemistry or organometallic pi-arene chemistry: strategies for inorganic crystal engineering.," J. Chem. Soc., Dalton Trans. p. 3855 (2000).

[28] C. Tuma, A. Boese, and N. Handy, "Predicting the binding energies of H-bonded complexes: a comparitive DFT study," Phys. Chem. Chem. Phys. **1,** 3939 (1999).

[29] N. A. Benedek, I. K. Snook, K. Latham, and I. Yarovsky, "Application of numerical basis sets to hydrogen bonded systems: a density functional theory study," J. Chem. Phys. **122,** 144102 (2005).

[30] J. B. Anderson, C. A. Traynor, and B. M. Boghosian, "An exact Quantum Monte Carlo calculation of the helium-helium intermolecular potential," J. Chem. Phys. **99,** 345 (1999).

[31] J. F. Dobson and J. Wang, "Successful test of a seamless van der Waals density functional," Phys. Rev. Lett. **82,** 2123 (1999).

[32] W. Kohn, Y. Meir, and D. E. Makarov, "van der Waals energies in density functional theory," Phys. Rev. Lett. **80,** 4153 (1998).

[33] A. G. Orpen, L. Brammer, F. H. Allen, O. Kennard, D. G. Watson, and R. Taylor, "Supplement. Tables of bond lengths determined by X-ray and neutron diffraction. Part 2. Organometallic compounds and co-ordination complexes of the d- and f-block metals," J. Chem. Soc., Dalton Trans. pp. S1–S83 (1989).

[34] M. A. V. Ribeiro da Silva, M. D. M. C. Ribeiro da Silva, M. J. S. Monte, J. M. Gonçalves, and E. M. R. Fernandes, "Energetics of metal-ligand binding in copper(II) and nickel (II) complexes of two Schiff bases," J. Chem. Soc., Dalton Trans. **7,** 1257 (1997).

[35] G. Al Takhin, H. A. Skinner, and A. A. Zaki, "Microcalorimetric studies on the thermal decomposition of platinum and palladium complexes containing phosphine or diene ligands," J. Chem. Soc., Dalton Trans. **3,** 371 (1984).

[36] B. S. Freiser, *Organometallic ion chemistry* (Kluwer, Dordrecht, 1996).

[37] H. A. Skinner and J. A. Connor, Pure Appl. Chem. **57,** 79 (1985).

[38] V. Coropceanu, M. Malagoli, D. A. da Silva Filho, N. E. Gruhn, T. G. Bill, and J. L. Bredàs, "Hole- and electron-vibrational couplings in oligoacene crystals: intramolecular contributions," Phys. Rev. Lett. **89,** 275503 (2002).

[39] J. E. Northrup, M. L. Tiago, and S. G. Louie, "Surface energetics and growth of pentacene," Phys. Rev. B **66,** 121404 (2002).

[40] C. Horn, I. Dance, D. Craig, M. Scudder, and G. Bowmaker, "A short but weak Cu–Cu interaction in $[Cu_2Br_5]^{2-}$, a crystal engineered $(Cu^{1.5+})_2$ confacial bitetrahedral complex," J. Am. Chem. Soc. **120,** 10549 (1998).

[41] W. L. McMillan, "Ground state of liquid He^4," Phys. Rev. **138,** A442 (1965).

[42] D. M. Ceperley and B. J. Alder, "Ground state of the electron gas by a stochastic method," Phys. Rev. Lett. **45,** 566 (1980).

[43] J. C. Grossman and L. Mitas, "Efficient Quantum Monte Carlo energies for molecular dynamics simulations," Phys. Rev. Lett. **94,** 056403 (2005).

[44] A. Aspuru-Guzik, O. El Akramine, J. Grossman, and W. A. Lester, "Quantum Monte Carlo for electronic excitations of free-base porphyrin," J. Chem. Phys. **120,** 3049 (2004).

[45] M. Born and J. R. Oppenheimer, "Zur Quantentheorie der Molekeln," Ann. Physik **84,** 457 (1927).

[46] D. R. Hartree, "The wave mechanics of an atom with non-coulombic central field: Parts I, II, III," Proc. Camb. Phil. Soc. **24,** 89,111,426 (1928).

[47] V. Fock, "Näherungsmethode zur Losung des quanten-mechanischen Mehrkörperprobleme," Z. Phys. **61,** 126 (1930).

[48] J. C. Slater, "The theory of complex spectra," Phys. Rev. **34,** 1293 (1929).

[49] W. Pauli, "Über den Zusammenhang des Abschlusses der Elecktronengruppen im Atom mit der Komplex Strucktur der Spektren," Z. Phys. **31,** 765 (1925).

[50] J. W. Strutt, *Theory of Sound, Vol. 1, Sec 88* (Reprint: Dover Publications, New York, 1945).

[51] W. Ritz, "Über eine methode zur losung gewisser variantionsprobleme der matahematischen physik," Reine Angew. Math. **135,** 1 (1908).

[52] L. H. Thomas, "The calculation of atomic fields," Proc. Camb. Phil. Soc. **23,** 542 (1927).

[53] E. Fermi, "Un Metodo Statistice per la Determinazionedi Alcune Proprieta dell' Atomo," Rend. Accad. Lincei. **6,** 602 (1927).

[54] P. Hohenberg and W. Kohn, "Inhomogenous electron gas," Phys. Rev. B **136,** 864 (1964).

[55] W. Kohn and L. J. Sham, "Self-consistent equations including exchange and correlation effects," Phys. Rev. A **140,** 1133 (1965).

[56] M. Levy, "Universal variational functionals of electron densities, first order density matrices and natural spin orbitals and solution of the v-representability problem," Proc. Natl. Acad. Sci. USA **76,** 6062 (1979).

[57] M. Levy, "Electron densities in search of hamiltonians," Phys. Rev. A **26,** 1200 (1982).

[58] E. Lieb, in *Physics as Natural Philosophy*, A. Shimony and H. Feshbach, eds., (MIT Press, Cambridge, 1982), p. 111.

[59] E. Lieb, "Density functionals for coulomb systems," Int. J. Quant. Chem. **24,** 243 (1983).

[60] A. Becke, "Density-functional exchange-energy approximation with correct asymptotic behavior," Phys. Rev. A **38,** 3098 (1988).

[61] J. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," Phys. Rev. Lett. **77,** 3865 (1996).

[62] R. W. Godby, M. Schlüter, and L. J. Sham, "Accurate exchange-correlation potential for silicon and its discontinuity on addition of an electron," Phys. Rev. Lett. **56,** 2415 (1986).

[63] J. P. Perdew and A. Zunger, "Self-interaction correction to density functional approximations for many-electron systems," Phys. Rev. B **23,** 5048 (1981).

[64] A. Boese, J. Martin, and N. Handy, "The role of the basis set: Assessing Density Functional Theory," J. Chem. Phys. **119,** 3005 (2003).

[65] R. Ditchfield, W. Hehre, and J. Pople, "Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules," J. Chem. Phys. **54,** 724 (1971).

[66] D. Woon and T. Dunning, "Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon," J. Chem. Phys. **98,** 1358 (1993).

[67] W. Herring, "A new method for calculating wave functions in crystals," Phys. Rev. **57,** 1169 (1940).

[68] R. Harrison, G. Fann, T. Yanai, and G. Beylkin, "Multiresolution quantum chemistry in multiwavelet bases," In *International Conference on Computational Science*, G. Goos, J. Hartmanis, and J. van Leeuwen, eds.,**4,** 103–110 (Springer-Verlag, Melbourne, Australia and St. Petersburg, Russia, 2003).

[69] W. J. Hehre, R. F. Stewart, and J. A. Pople, "Self-consistent molecular-orbital methods. I. Use of Gaussian expansions of Slater-type atomic orbitals," J. Chem. Phys. **51,** 2657 (1969).

[70] G. Wannier, "Dynamics of band electrons in electric and magnetic fields," Rev. Mod. Phys. **34,** 645 (1962).

[71] B. Delley, "An all-electron numerical method for solving the local density functional for polyatomic molecules," J. Chem. Phys. **92,** 508 (1990).

[72] B. Delley, "From molecules to solids with the Dmol^3 approach," J. Chem. Phys. **113,** 7756 (2000).

[73] M. Meunier, N. Quirke, and D. Binesti, "The calculation of the electron affinity of atoms and molecules," Mol. Sim. **23,** 109 (1999).

[74] C. R. A. Catlow, F. Cora, and A. A. Sokol, "Electron spin localisation and correlation effects for point defects in semi-ionic solids," Comp. Mat. Sci. **17,** 312 (2000).

[75] S. B. Legoas, D. S. Galvao, V. Rodrigues, and D. Ugarte, "Origin of anomalously long interatomic distances in suspended gold chains," Phys. Rev. Lett. **88,** 076105 (2002).

[76] Y. J. Zhao, W. T. Geng, A. J. Freeman, and B. Delley, "Structural, electronic and magnetic properties of alpha- and beta-MnAs: LDA and GGA investigations," Phys. Rev. B **65,** 113202 (2002).

[77] G. S. Tschumper, M. D. Kelty, and H. F. Schaefer, "Subtle basis set effects on hydrogen bonded systems," Mol. Phys. **96,** 493 (1999).

[78] A. Halkier, H. Koch, P. Jorgensen, O. Christiansen, I. M. Beck Nielsen, and T. Helgaker, "A systematic ab initio study of the water dimer in hierarchies of basis sets and correlation models," Theor. Chem. Acc. **97,** 150 (1997).

[79] A. Halkier, W. Klopper, T. Helgaker, P. Jorgensen, and P. R. Taylor, "Basis set convergence of the interaction energy of hydrogen-bonded complexes," J. Chem. Phys. **111,** 9157 (1999).

[80] M. Schutz, S. Brdarski, P. O. Widmark, R. Lindh, and G. Karlstrom, "The water dimer interaction energy: convergence to the basis set limit at the correlated level," J. Chem. Phys. **107,** 4597 (1997).

[81] J. E. Del Bene, W. B. Person, and K. Szczepaniak, "Properties of hydrogen-bonded complexed obtained from the B3LYP functional with 6-31G(d,p) and 6-31+G(d,p) basis sets: comparison with MP2/6-31+G(d,p) results and experimental data," J. Phys. Chem. **99,** 10705 (1995).

[82] J. Altmann, N. Handy, and V. Ingamells, "A study of the performance of numerical basis sets in DFT calculations on sulfur-containing molecules," Int. J. Quant. Chem. **57,** 533–542 (1996).

[83] *DMol*^{3} *User Guide* (Accelrys Inc., San Diego, 2003).

[84] C. Lee, W. Yang, and R. Parr, "Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density," Phys. Rev. B **37,** 785 (1988).

[85] B. Miehlich, A. Savin, H. Stoll, and H. Preuss, Chem. Phys. Lett. **157,** 200 (1989).

[86] F. Hamprecht, A. Cohen, D. Tozer, and N. Handy, "Development and assessment of new exchange-correlation functionals," J. Chem. Phys. **109,** 6264 (1998).

[87] M. J. Frisch *et al.*, "Gaussian03,", (2003).

[88] T. Dunning, "Gaussian basis functions for use in molecular calculations. III. Contraction of (10s6p) atomic basis sets for the first-row atoms," J. Chem. Phys. **55,** 716 (1971).

[89] P. Pulay, "Improved SCF convergence acceleration," J. Comp. Chem. **3,** 556 (1982).

[90] S. Boys and F. Bernardi, "The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors.," Mol. Phys. **19,** 553 (1970).

[91] K. Peterson and T. Dunning, "Benchmark calculations with correlated molecular wave functions. VII. Binding energy and structure of the HF dimer," J. Chem. Phys. **102,** 2032 (1995).

[92] K. Hermansson and M. Probst, "Correlation between intramolecular bond distances and stretching vibrations for polar molecules: an ab initio study," Int. J. Quant. Chem. **63,** 537 (1997).

[93] S. I. Lu, "A diffusion quantum Monte Carlo study of geometries and harmonic frequencies of molecules," J. Chem. Phys. **120,** 14 (2004).

[94] K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure, IV. Constants of Diatomic Molecules* (Van Nostrand Reinhold, New York, 1979).

[95] W. S. Benedict, N. Gailar, and E. K. Plyler, J. Chem. Phys. **24,** 1139 (1956).

[96] W. Koch and M. Holthausen, *A Chemist's Guide to Density Functional Theory*, 2nd ed. (Wiley-VCH, Weinheim, 2002).

[97] H. Ni, J. M. Serafin, and J. J. Valentini, "Dynamics of the vibrational predissociation of HCl dimer," J. Chem. Phys. **113,** 3055 (2000).

[98] W. Klopper and H. Luthi, "The MP2 limit correction applied to coupled cluster calculations of the electronic dissociation energies of the hydrogen fluoride and water dimers," Mol. Phys. **96,** 559 (1999).

[99] C. Maerker, P. von R. Schleyer, K. R. Liedl, M. Quack, and M. A. Suhm, "A critical analysis of electronic density functionals for structural, energetic, dynamic and magnetic properties of hydrogen fluoride clusters," J. Comp. Chem. **18,** 1695 (1997).

[100] E. J. Bohac, M. D. Marshall, and R. E. Miller, "Initial state effects in the vibrational predissociation of hydrogen fluoride dimer," J. Chem. Phys. **96,** 6681 (1992).

[101] A. S. Pine and B. J. Howard, "Hydrogen bond energies of the HF and HCl dimers from absolute infrared intensities," J. Chem. Phys. **84,** 590 (1986).

[102] M. Quack and M. A. Suhm, "Potential energy surfaces, quasiadiabatic channels, rovibrational spectra, and intramolecular dynamics of $(HF)_2$ and its isotopomers from quantum Monte Carlo calculations," J. Chem. Phys. **95,** 28 (1991).

[103] B. J. Howard, T. R. Dyke, and W. Klemperer, "The molecular beam spectrum and the structure of the hydrogen fluoride dimer," J. Chem. Phys. **81,** 5417 (1984).

[104] H. S. Gutowsky, C. Chuang, J. D. Keen, T. D. Klots, and T. Emilsson, "Microwave rotational spectra, hyperfine interactions, and structure of the hydrogen fluoride dimers," J. Chem. Phys. **83,** 2070 (1985).

[105] W. Klopper, J. van Duijneveldt-van de Rijdt, and F. van Duijneveldt, "Computational determination of equilibrium geometry and dissocation energy of the water dimer," Phys. Chem. Chem. Phys. **2,** 2227 (2000).

[106] L. A. Curtiss, D. J. Frurip, and M. Blander, "Studies of molecular association in H_2O and D_2O vapors by measurement of thermal conductivity," J. Chem. Phys. **71,** 2703 (1979).

[107] J. Reimers, R. Watts, and M. Klein, "Intermolecular potential functions and the properties of water," Chem. Phys. **64,** 95 (1982).

[108] M. Elrod and R. Saykally, "Determination of the intermolecular potential energy surface for (HCl)2 from vibrationrotationtunneling spectra," J. Chem. Phys. **103,** 933 (1995).

[109] L. Gonzalez, O. Mo, M. Yanez, and J. Elguero, "Very strong hydrogen bonds in neutral molecules: The phosphinic acid dimers," J. Chem. Phys. **109,** 2685 (1998).

[110] *CRC Handbook of Tables for Mathematics*, 4th ed., R. C. Weast and S. M. Selby, eds., (The Chemical Rubber Co., Cleveland, 1970).

[111] M. Meot-Ner, "The ionic hydrogen bond," Chem. Rev. **105,** 213 (2005).

[112] G. Gronert, "Theoretical studies of proton transfers. 1. The potential energy surfaces of the identity reactions of the first- and second-row non-metal hydrides with their conjugate bases," J. Am. Chem. Soc. **115,** 10258 (1993).

[113] J. R. Pliego and J. M. Riveros, "Ab initio study of the hydroxide ion-water clusters," J. Chem. Phys. **112,** 4045 (2000).

[114] H. M. Lee, P. Tarkeshwar, and K. S. Kim, "Structures, energetics and spectra of hydrated hydroxide anion clusters," J. Chem. Phys. **121,** 4657 (2004).

[115] M. D. Newton and S. Ehrenson, "Ab initio studies on the structures and energetics of inner and outer-shell hydrates of the proton and hydroxide ion," J. Am. Chem. Soc. **93,** 4971 (1971).

[116] J. E. Del Bene, "Ab initio molecular orbital study of the structures and energies of neutral and charged bimoleculear complexes of H_2O with the hydrides AH_n (A = N, O, F, P, S and Cl)," J. Phys. Chem. **92,** 2874 (1988).

[117] S. Wolfe, S. Hoz, C. K. Kim, and K. Yang, "Barrier widths, barrier heights and the origins of anomalous kinetic H/D isotope effects," J. Am. Chem. Soc. **112,** 4186 (1990).

[118] M. Meot-Ner and L. W. Sieck, "The ionic hydrogen bond and ion solvation. 5. $OH\cdots O^-$ bonds. Gas-phase solvation and clustering of alkoxide and carboxylate anions.," J. Am. Chem. Soc. **108,** 7525 (1986).

[119] K. Latham, A. M. Coyle, C. J. Rix, A. Fowless, and J. M. White, Polyhedron .

[120] K. Nagarajan, K. P. Shelly, R. R. Perkins, and R. Stewart, "Arylphosphonic acids. I. Substituent effects on their first and second dissociations," Can. J. Chem. **65,** 1729 (1987).

[121] A. I. Kitaigorodsky, *Molecular Crystals and Molecules* (Academic Press, New York, 1973).

[122] A. Martin and A. G. Orpen, "Structural Systematics. 6. Apparent flexibility of metal complexes in crystals," J. Am. Chem. Soc. **118,** 1464 (1996).

[123] B. Ali, I. Dance, M. Scudder, and D. Craig, "Dimorphs of $(Ph_4P)_2[Cd_2(SPh)_6]$: Crystal packing analyses and the interplay of intermolecular and intramolecular energies," Cryst. Growth Des. **2,** 601 (2002).

[124] C. Ceccarelli, G. A. Jeffrey, and R. Taylor, "A survey of O–H$\cdots$O hydrogen bond geometries determined by neutron diffraction," J. Mol. Struct. **70,** 255 (1981).

[125] F. A. Cotton, *Advanced Inorganic Chemistry*, 6th ed. (Wiley, New York, 1999).

[126] R. S. Mulliken, "Molecular compounds and their spectra. II," J. Am. Chem. Soc. **74,** 811 (1952).

[127] K. Morokuma, "Why do molecules interact? The origin of electron donor-acceptor complexes, hydrogen bonding, and proton affinity," Acc. Chem. Res. **10,** 294 (1977).

[128] S. G. Wang and W. H. E. Schwarz, "Density functional study of first row transition metal dihalides," J. Chem. Phys. **109,** 7252 (1998).

[129] E. Ruiz, D. R. Salahub, and A. Vela, "Charge-transfer complexes: Stringent tests for widely used density functionals," J. Phys. Chem. **100,** 12265 (1996).

[130] T. Bally and G. N. Sastry, "Incorrect dissociation behaviour of radical ions in density functional calculations," J. Phys. Chem. A **101,** 7923 (1997).

[131] J. Gräfenstein, E. Kraka, and D. Cremer, "The impact of the self-interaction error on the density functional theory description of dissociating radical cations: Ionic and covalent dissociation limits," J. Chem. Phys. **120,** 524 (2004).

[132] H. Chermette, I. Ciofini, F. Mariotti, and C. Daul, "Correct dissociation behaviour of radical ions such as $H_2{}^+$ in density functional calculations," J. Chem. Phys. **114,** 1447 (2001).

[133] A. D. Becke, "A real-space model of nondynamical correlation," J. Chem. Phys. **119,** 2972 (2003).

[134] A. D. Becke, "Real-space post-Hartree-Fock correlation models," J. Chem. Phys. **122,** 064101 (2005).

[135] U. El-Ayaan, A. Paulovicova, S. Yamada, and Y. Fukuda, "The crystal structure of bis[*N*-(2,6-diisopropylphenyl)imino] acenaphthene and studies of its copper(I) and copper(II) complexes," J. Coord. Chem. **56,** 373 (2003).

[136] M. T. Miller, P. K. Gantzel, and T. B. Karpishin, "Structures of the Copper(I) and Copper (II) complexes of 2,9-diphenyl-1,10-phenanthroline: Implications for excited state structural distortion," Inorg. Chem. **37,** 2285 (1998).

[137] F. K. Klemens, C. E. A. Palmer, S. M. Rolland, P. E. Fanwick, D. R. McMillin, and J. P. Savauge, "The crystal and molecular structures of the dpp ligand and $[Cu(dpp)_2][CuCl_2]\cdot CH_3CN$," New J. Chem. **14,** 129 (1990).

[138] M. Ruthkosky, F. N. Castellano, and G. J. Meyer, "Photodriven electron and energy transfer from copper phenanthroline excited states," Inorg. Chem. **35,** 6406 (1996).

[139] H. Chermette, I. Ciofini, F. Mariotti, and C. Daul, "A posteriori corrections to systematic failures of standard density functionals: The dissociation of two-center three-electron systems," J. Chem. Phys. **115,** 11068 (2001).

[140] Y. Zhang and W. Yang, "A challenge for density functionals: self-interaction error increases for systems with a non-integer number of electrons," J. Chem. Phys. **109,** 2604 (1998).

[141] J. G. Àngyàn, (2005).

[142] B. Delley and P. Scherrer, "DMol, a standard tool for density functional calculations: Review and Advances," in *Modern Density Functional Theory: A Tool for Chemists*, J. M. Seminario and P. Politzer, eds., (Elsevier, Amsterdam, 1995), Vol. 2.

[143] G. Ortiz, M. Harris, and P. Ballone, "Correlation energy, structure factor, radial density distribution function and momentum distribution of the spin-polarized electron gas," Phys. Rev. B **50,** 1391 (1994).

[144] R. Q. Hood, M. Y. Chou, A. J. Williamson, G. Rajagopal, R. J. Needs, and W. M. C. Foulkes, "Quantum Monte Carlo investigation of exchange and correlation in silicon," Phys. Rev. Lett. **78,** 3350 (1997).

[145] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," J. Chem. Phys. **21,** 1087 (1953).

[146] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," Rev. Mod. Phys. **73,** 33 (2001).

[147] P. Reynolds, D. Ceperley, B. Alder, and W. Lester, "Fixed-node quantum Monte Carlo for molecules," J. Chem. Phys. **77,** 5593 (1982).

[148] J. B. Anderson, "Quantum chemistry by random walk," J. Chem. Phys. **65,** 4121 (1976).

[149] R. J. Jastrow, "Many-body problem with strong forces," Phys. Rev. **98,** 1479 (1955).

[150] N. D. Drummond, M. D. Towler, and R. J. Needs, "Jastrow correlation factor for atoms, molecules and solids," Phys. Rev. B **70,** 235119 (2004).

[151] N. D. Drummond and R. J. Needs, "Variance-minimization scheme for optimizing Jastrow factors," Phys. Rev. B **72,** 085124 (2005).

[152] A. Lüchow, J. B. Anderson, and D. Feller, "Improved estimates of the total correlation energy in the ground state of the water molecule," J. Chem. Phys. **106,** 7706 (1997).

[153] J. A. Odutola and T. R. Dyke, "Partially deuterated water dimers: Microwave spectra and structure," J. Chem. Phys. **72,** 5062 (1980).

[154] J. G. C. M. van Duijneveldt-van de Rijdt and F. B. van Duijneveldt, "Convergence to the basis set limit in ab initio calculations at the correlated level on the water dimer," J. Chem. Phys. **97,** 5019 (1992).

[155] E. M. Mas and K. Szalewicz, "Effects of monomer geometry and basis set saturation on computed depth of water dimer potential," J. Chem. Phys. **104,** 7606 (1996).

[156] P. O. Åstrand, G. Karlström, A. Engdahl, and B. Nelander, "Novel model for calculating the intermolecular part of the infrared spectrum for molecular complexes," J. Chem. Phys. **102,** 3534 (1995).

[157] P. O. Widmark, P. A. Malmqvist, and B. O. Roos, "Density matrix averaged atomic natural orbital (ANO) basis sets for correlated molecular wavefunctions," Theo. Chim. Acta **77,** 291 (1990).

[158] R. J. Needs, M. D. Towler, N. D. Drummond, and P. R. C. Kent, "CASINO version 1.7.1 User Manual,", (2004).

[159] W. Ermler and C. Kern, "Simple procedure for estimating the Hartree-Fock limit energies of molecules," J. Chem. Phys. **61,** 3860 (1974).

[160] S. Manten and A. Lüchow, "On the accuracy of the fixed-node diffusion Quantum Monte Carlo method," J. Chem. Phys. **115,** 5362 (2001).

[161] A. Lüchow and R. F. Fink, "On the systematic improvement of fixed-node diffusion Quantum Monte Carlo energies using pair natural orbital CI guide wavefunctions," J. Chem. Phys. **113,** 8457 (2000).

[162] H. Müller, W. Kutzelnigg, and J. Noga, "A CCSD(T)-R12 study of the ten-electron systems Ne, F^-, HF, H_2O, NH_3, NH_4^+, and CH_4," Mol. Phys. **92,** 535 (1997).

[163] D. Feller, C. Boyle, and E. Davidson, "One-electron properties of several small molecules using near Hartree-Fock limit basis sets," J. Chem. Phys. **86,** 3424 (1987).

[164] E. M. Mas, R. Bukowski, K. Szalewicz, G. C. Groenenboom, P. E. S. Wormer, and A. van der Avoird, "Water pair potential of near spectroscopic accuracy. I. Analysis of potential surface and virial co-efficients," J. Chem. Phys. **113,** 6687 (2000).

[165] C. Diedrich, A. Lüchow, and S. Grimme, "Weak intermolecular interactions calculated with diffusion Monte Carlo," J. Chem. Phys. **123,** 184106 (2005).

[166] J. Grossman, "Benchmark Quantum Monte Carlo calculations," J. Chem. Phys. **117,** 1434 (2002).

[167] L. Wagner and L. Mitas, "A Quantum Monte Carlo study of electron correlation in transition metal oxygen molecules," Chem. Phys. Lett. **370,** 412 (2003).

[168] D. Bressanini, G. Morosi, and S. Tarasco, "An investigation of nodal structures and the construction of trial wave functions," J. Chem. Phys. **123,** 204109 (2005).

[169] J. R. Trail and R. J. Needs, "Smooth relativistic Hartree-Fock pseudopotentials for H to Ba and Lu to Hg," J. Chem. Phys. **122,** 174109 (2005).

[170] J. R. Trail and R. J. Needs, "Norm-conserving Hartree-Fock pseudopotentials and their asymptotic behaviour," J. Chem. Phys. **122,** 014112 (2005).